Serial No. 09/538,396
Group Art Unit: 1638

REMARKS

Reconsideration of the present application is respectfully requested. Claims 2-8, and 12-15 are pending. Claims 9-11 have been cancelled as belonging to a non-elected invention. The right to pursue these claims in a continuing application is reserved. No change of inventorship is necessary. Claim 1 has been cancelled and rewritten as new claims 12-15. Claims 2-4 have been amended to correct dependency. Support for these claims is found in the claims as originally filed, and throughout the specification. No new matter has been added.

Applicant has amended the specification to delete references to Internet hyperlinks.

The marked up version of these amendments is found on a separate sheet attached to this amendment and titled **"Version with Markings to Show Changes Made."** It is respectfully requested that the amendments be entered.

Election/Restriction

The Examiner has issued a restriction requirement, and has required election of either the invention of Group I (Claims 1-8) or Group II (Claims 9-11). Applicants hereby affirm the election to prosecute the claims of Group I, with traverse as filed 8/31/01. Applicants expressly reserve the right to file a divisional applications relating to and claiming the inventions of Group II and/or Group III. No change of inventorship is required due to this election of Group I.

Rejections under 35 U.S.C. §101:

Claims 1-8 are rejected under 35 U.S.C. §101 as not having either a credible asserted utility or a well-established utility. Claim 1 has been cancelled and rewritten as claims 12-15, the rejection will be discussed as it applies to these claims.

Serial No. 09/538,396
Group Art Unit: 1638

The Examiner asserts that "No function of said polynucleotides are recited."

Applicants have rewritten claim 1 as new claims 12-15. New claims 12 and 13 now recite "wherein the polynucleotide encodes a polypeptide with Rad50 activity." New claim 15 is dependent on new claim 12, so also requires that 30 contiguous nucleotides come from a polynucleotide which encodes a polypeptide with Rad50 activity. New claim 14 is directed to polynucleotides which hybridize to SEQ ID NO: 1. Therefore, new claims 12-14 and dependent claims 2-8, and 15 do recite the function of the polynucleotides.

The Examiner asserts "Applicants assert that a polynucleotide having 80% sequence identity to SEQ ID NO: 1 would have Rad50 activity. However it is unclear what would be the utility of said polynucleotide if the 20% lack of identity falls in a region crucial for the Rad50 activity."

Applicants have rewritten claim 1 as new claims 12-15. This rejection will be discussed as it applies to new claim 12. In the preamble, Claim 12 recites "An isolated polynucleotide encoding a polypeptide with Rad50 activity ". Therefore, only polynucleotides with 80% sequence identity to SEQ ID NO: 1, which also encode a polypeptide with Rad50 activity are claimed. Further, not all embodiments must have utility for the invention as a whole to have utility. Inoperable embodiments of the claimed invention do not eliminate the utility of the operable embodiments. As it is stated in the MPEP 2107 II, page 2100-25: "... as the Federal Circuit has stated, '[t]o violate [35 U.S.C.] 101 the claimed device must be totally incapable of achieving a useful result.' Brooktree Corp. v. Advanced Micro Devices, Inc., 977 F.2d 1555, 1571, 24 USPQ2d 1401, 1412 (Fed. Cir. 1992)".

The Examiner states "No data that relates SEQ ID NO: 1 or SEQ ID NO: 2 to Rad50 activity has been shown."

Applicants respectfully disagree, page 1, line 15 – page 2, line 15 of the specification clearly details the well-established activity and features of Rad50 polypeptides. Rad50 has been shown to be involved in DNA recombination and

9

Serial No. 09/538,396
Group Art Unit: 1638

repair, the present invention proposes to use the well established activity of Rad50 to improve transformation efficiency in plants, therefore establishing specific and substantial utility for the present invention. Page 2, lines 18-25, and in Example 4 on pages 62-64, of the specification discuss the structural features shared by SEQ ID NO:2 of the present invention and other known Rad50 proteins, including the predicted molecular weight, the presence of two ATP-binding sites (Walker boxes), nuclear localization signals, heptad repeats, and leucine zippers. In Appendix A, Applicants submit a multiple sequence alignment of SEQ ID NO: 2 with several other Rad50 proteins. Identical and conserved amino acids, relative to SEQ ID NO: 2, are highlighted. The multiple sequence alignment illustrates the extensive homology, over the entire length of SEQ ID NO: 2, to other Rad50 proteins.

The Examiner asserts "However, the state of the art as exemplified by Bork et al suggests that a 31.7% of sequence identity of Applicant's SEQ ID NO: 2 with the known protein is insufficient to predictably determine the function of Applicant's protein."

The identification of SEQ ID NO: 1 and SEQ ID NO: 2 as Rad50 polynucleotide and polypeptide respectively, is not based merely on percent sequence identity alone, but is based on an analysis of several features, such as molecular weight, and sequence homology to known conserved domains contained in Rad50. These features include the presence and positioning of ATP-binding sites, nuclear localization signals, heptad repeats, and leucine zippers. As illustrated in the multiple sequence alignment presented in Appendix A, there is substantial homology to other Rad50 proteins across the entire length of SEQ ID NO: 2. Therefore, the Applicant has established a credible utility for the sequences of the present invention.

While Bork (Genome Research 10:398-400, 2000) certainly wishes to warn about the potential limits to extrapolating the data of high-throughput technologies which automatically annotate genomic sequencing efforts, he does not state that

Serial No. 09/538,396
Group Art Unit: 1638

computer-based homology searches are invalid or questionable. In fact, on page 400, second column, second paragraph Bork states " However there is still no doubt that sequence analysis is extremely powerful and that the generation of hypotheses derived by computational methods will be more and more often the first successful step in the design of experiments. If 70% of such experiments were successful, the speed of scientific discoveries would grow exponentially."

The Applicants also respectfully draw the Examiner's attention to the Utility Examination Guidelines, Official Gazette, January 30, 2001 which state "... when a patent application claiming a nucleic acid asserts a specific, substantial, and credible utility, and bases the assertion upon homology to existing nucleic acids having an accepted utility, the asserted utility must be accepted by the examiner unless the Office has sufficient evidence or sound scientific reasoning to rebut such an assertion." The Guidelines further state "[A] 'rigorous correlation' need not be shown in order to establish practical utility; 'reasonable correlation' is sufficient." *Fujikawa v. Wattanasin*, 93 F.3d 1559, 1565, 39 USPQ2d 1895, 1900 (Fed. Cir. 1996).

The Examiner cites Lazar et al. (*Mol Cell Biol* 1998 8(3):1247-1252), and Broun et al. (*Science* 1998 282:131-133), which provide examples of very specific limited amino acid changes which result in elimination or alteration of the experimental protein's catalytic activity.

There are usually many positions within the primary sequence of a protein where substitution has little or no effect on the protein's activity, there are even cases where these sites are also part of a binding domain or active site. There are even cases where substitution of a particular amino acid can increase catalytic activity. As was stated earlier, the invention is directed to compositions of Rad50 and its activities, non-functional embodiments are not claimed and do not eliminate the utility of the function embodiments set forth in the claims.

Applicants believe that the present invention has a well-established utility for which they have proposed specific, substantial and credible uses in the present

11

Serial No. 09/538,396
Group Art Unit: 1638

application. Applicants have properly addressed by argument and amendment the grounds for the rejection of originally filed claims 1-18 under 35 U.S.C. §101 as it would apply to pending claims 2-8, and 12-15, and respectfully request that the rejection of the claims under 35 U.S.C §101 be withdrawn.

Rejections under 35 U.S.C. §112, first paragraph – Utility:

As the Applicants have responded to the utility rejection under 35 U.S.C. §101, the concomitant rejection of claims 1-8 under 35 U.S.C. §112, first paragraph based on a lack of utility should be withdrawn and not applied to pending claims 2-8, and 12-15.

Rejections under 35 U.S.C. §112, first paragraph – Written Description:

Claims 1-8 are rejected under 35 U.S.C. §112, first paragraph as containing subject matter which was not described in the specification in such a way as to reasonably convey to one skilled in the relevant art that the inventors, at the time the application was filed, had possession of the claimed invention. This rejection will be discussed as it pertains to original claims 2-8, and new claims 12-15.

The Examiner states: "Claim 1, part (c) is drawn to a polynucleotide having sequence amplified from a Zea mays nucleic acid library. No specific chemical or physical characteristics were disclosed for other polynucleotide sequences having sequence amplified from a Zea mays nucleic acid library. The claim encompasses undiscovered genes and undisclosed regions of Zea mays nucleic acid library outside of SEQ ID NO: 1 which applicant is not in possession of at the time of filing."

Claim 1 was cancelled. Original claim 1, part (c) is presented as new claim 13. The rejection will be discussed as it may be applied to new claim 13.

Claim 13 claims "A polynucleotide amplified from a *Zea mays* nucleic acid library using primers which selectively hybridize, under stringent hybridization

Serial No. 09/538,396
Group Art Unit: 1638

conditions, to loci within a polynucleotide of SEQ ID NO: 1, wherein the polynucleotide encodes a polypeptide with Rad50 activity."

Applicants respectfully disagree that no specific chemical and physical characteristics are disclosed. The chemical and physical characteristics include the disclosure of the full-length sequence of SEQ ID NO: 1, and the limitations that the polynucleotide be amplified from a Zea mays nucleic acid library, the primers used must selectively hybridize under stringent conditions, the primers must hybridize to loci within SEQ ID NO: 1. Claim 13 also states the amplified polynucleotide must encode a polypeptide with Rad50 activity.

Applicants clearly define amplified on page 4, lines 10-12; selectively hybridizes on page 13, lines 3-9; and stringent hybridization conditions on page 13, line 30 – page 15, line 16. Applicants provide guidance regarding amplification of polynucleotides on page 24, line 15 – page 26, line 10 and page 35, line 29 – page 36, line 19; construction of nucleic acid libraries on page 32, line 11 – page 35, line 9, and Example 1 on pages 59-60. Claim 13 clearly claims the amplification primers used must selectively hybridize under stringent conditions to loci within SEQ ID NO: 1.

Given the disclosure of a full-length maize Rad50 polynucleotide in SEQ ID NO: 1, guidance on amplification and nucleic acid library construction, and the clearly defined parameters of claim 13, the subject matter of claims 2-8, and 12-15 was reasonably conveyed to one of skill in the art and indicated the Applicants had possession of the claimed invention at the time of filing. Therefore, it is respectfully requested that the rejection of claims under 35 U.S.C. §112, first paragraph be withdrawn.

Rejections under 35 U.S.C. § 102:

Claims 1-8 have been rejected under 35 U.S.C. § 102(b) as being anticipated by Boudet et al. (US Patent 5,451,514).

Serial No. 09/538,396
Group Art Unit: 1638

The Examiner asserts "The claims read on a polynucleotide with 2-bases, since any two bases would hybridize and would be complementary to the claimed polynucleotide."

Claim 1 was cancelled and rewritten as new claims 12-15. Original claim 1, part (d) is now presented as claim 14. Original claim 1, part (f) is not presented as claim 12, part (d). The rejection will be addressed as it may apply to these new claims.

The Applicants respectfully disagree that the claims encompass 2 nucleotide fragments. Sequences of only two nucleotides in length would not even **anneal** to the nucleic acid of the present invention under most conditions, much less **selectively hybridize** to the nucleic acid of the present invention as it is defined on page 13, lines 3 - 9 under **stringent conditions** as described on pages 13, line 30 – page 15, line 16 of the specification. Using the quick calculation for melting temperature (Tm) of 4° C for every G or C nucleotide, or 2° C for every A or T nucleotide (Wallace formula), one can quickly calculate the approximate maximum Tm for a two nucleotide sequence to be 8° C, annealing temperature is generally calculated as 5° C lower than the Tm, or 3° C in this case. It is apparent that subsequences of only 2 nucleotides in length are not capable of annealing to, much less selectively hybridizing with, the nucleic acid of the present invention, therefore the rejection of claim 1 (d) and (f) should be withdrawn and not applied to new claims 12 and 14.

The Applicants respectfully traverse the rejection under 35 U.S.C. § 102(b). As it is stated in the MPEP 2131 page 2100-54 "To anticipate a claim, the reference must teach every element of the claim. 'A claim is anticipated only if each and every element as set forth in the claim is found, either expressly or inherently described, in a single prior art reference.'"

Boudet et al do not disclose a polynucleotide which encodes a polypeptide with Rad50 activity, or a polynucleotide that selectively hybridizes to SEQ ID NO: 1,

14
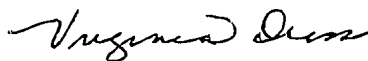
Serial No. 09/538,396
Group Art Unit: 1638

or a polynucleotide which is fully complementary to a polynucleotide which encodes
a polypeptide with Rad50 activity. Therefore, Boudet et al does not anticipate the
claims and the rejection under 35 U.S.C. § 102(b) should be withdrawn.

CONCLUSION

In light of the foregoing remarks and amendments, withdrawal of the
outstanding rejections and allowance of all of the remaining claims is respectfully
requested.

Respectfully submitted,

Virginia Dress
Agent for Applicant(s)
Registration No. 48,243

PIONEER HI-BRED INTERNATIONAL, INC.
Corporate Intellectual Property
7100 N.W. 62nd Avenue
P.O. Box 1000
Johnston, Iowa 50131-1000
Phone: (515) 270-4192
Facsimile: (515) 334-6883

Serial No. 09/538,396
Group Art Unit: 1638

VERSION WITH MARKINGS TO SHOW CHANGES MADE

The Applicants have used underlining to denote additions to the original text and square brackets [ ] to denote deletions of the original text.

In the Title:

The title found on the cover page has been amended as follows:

[Maize] Rad50 Orthologue and Uses Thereof

In the Specification:

Paragraph beginning at line 18 of page 2 has been amended as follows:

The present invention describes the maize Rad50 protein, which clearly possesses features characteristic of other Rad50 proteins, and has a calculated molecular weight of ~152.5 kDa. The maize Rad50 protein is characterized by the presence of an ATP binding site in the N-terminal region, a second nucleotide binding site in the C-terminal region, putative nuclear localization signals, and heptad-repeats. The presence of extensive leucine zipper structures appears to be another striking feature of the Rad50 proteins. These are also found in the maize Rad50 protein and are indicated in **bold** in [Figure 1] Example 4. The present invention also describes a maize Rad50 polynucleotide sequence. The maize Rad50 orthologue of the present invention was used as a probe to map the maize RAD50 gene(s) to the short arm of chromosome 4.

16

Serial No. 09/538,396
Group Art Unit: 1638

Paragraph beginning at line 9 of page 17 has been amended as follows:

Software for performing BLAST analyses is publicly available, e.g., through the National Center for Biotechnology Information [(http://www.ncbi.nlm.nih.gov/)]. This algorithm involves first identifying high scoring sequence pairs (HSPs) by identifying short words of length W in the query sequence, which either match or satisfy some positive-valued threshold score T when aligned with a word of the same length in a database sequence. T is referred to as the neighborhood word score threshold. These initial neighborhood word hits act as seeds for initiating searches to find longer HSPs containing them. The word hits are then extended in both directions along each sequence for as far as the cumulative alignment score can be increased. Cumulative scores are calculated using, for nucleotide sequences, the parameters M (reward score for a pair of matching residues; always > 0) and N (penalty score for mismatching residues; always < 0). For amino acid sequences, a scoring matrix is used to calculate the cumulative score. Extension of the word hits in each direction are halted when: the cumulative alignment score falls off by the quantity X from its maximum achieved value; the cumulative score goes to zero or below, due to the accumulation of one or more negative-scoring residue alignments; or the end of either sequence is reached. The BLAST algorithm parameters W, T, and X determine the sensitivity and speed of the alignment. The BLASTN program (for nucleotide sequences) uses as defaults a wordlength (W) of 11, an expectation (E) of 10, a cutoff of 100, M=5, N=-4, and a comparison of both strands. For amino acid sequences, the BLASTP program uses as defaults a wordlength (W) of 3, an expectation (E) of 10, and the BLOSUM62 scoring matrix (see Henikoff & Henikoff (1989) Proc. Natl. Acad. Sci. USA 89:10915).

17

Serial No. 09/538,396
Group Art Unit: 1638

Paragraph beginning at line 8 of page 62 has been amended as follows:

Gene identities were determined by conducting BLAST (Basic Local
Alignment Search Tool; Altschul, S. F., et al., (1990) J. Mol. Biol. 215:403-410[; see
also www.ncbi.nlm.nih.gov/BLAST/]) searches under default parameters for
similarity to sequences contained in the BLAST "nr" database (comprising all non-
redundant GenBank CDS translations, sequences derived from the 3-dimensional
structure Brookhaven Protein Data Bank, the last major release of the SWISS-PROT
protein sequence database, EMBL, and DDBJ databases). The cDNA sequences
were analyzed for similarity to all publicly available DNA sequences contained in the
"nr" database using the BLASTN algorithm. The DNA sequences were translated in
all reading frames and compared for similarity to all publicly available protein
sequences contained in the "nr" database using the BLASTX algorithm (Gish, W.
and States, D. J. Nature Genetics 3:266-272 (1993)) provided by the NCBI. In some
cases, the sequencing data from two or more clones containing overlapping
segments of DNA were used to construct contiguous DNA sequences.

The Abstract beginning at line 1 of page 67 has been amended as follows:

ABSTRACT OF THE DISCLOSURE
The invention provides isolated [maize] Rad50 nucleic acids and their encoded
proteins. The present invention provides methods and compositions relating to
altering Rad50 levels in plants. The invention further provides recombinant
expression cassettes, host cells, transgenic plants, and antibody compositions.

In the Claims:

Claims 1 and 9-11 have been cancelled without prejudice.

Serial No. 09/538,396
Group Art Unit: 1638

Claims 2, 3 and 4 have been amended as follows:

2.   (Amended) A recombinant expression cassette, comprising a member of
        claim [1] 12 operably linked[, in sense or anti-sense orientation,] to a
        promoter.

3.   (Amended) A host cell comprising a polynucleotide of claim [2] 12.

4.   (Amended) A transgenic plant comprising a recombinant expression cassette
        [of claim 2] comprising a polynucleotide of claim 12.

New claims 12-15 have been added as follows:

12.   An isolated polynucleotide encoding a polypeptide with Rad50 activity
        comprising a polynucleotide selected from the group consisting of:
        (a)   a polynucleotide having at least 80% sequence identity over the entire
                length of the reference sequence, as determined by the GAP program
                under default parameters, to a polynucleotide of SEQ ID NO: 1;
        (b)   a polynucleotide encoding a polypeptide of SEQ ID NO: 2;
        (c)   a polynucleotide of SEQ ID NO: 1;
        (d)   a polynucleotide which is fully complementary to a polynucleotide of
                (a), (b), or (c).

13.   A polynucleotide amplified from a Zea mays nucleic acid library using primers
        which selectively hybridize, under stringent hybridization conditions, to loci
        within a polynucleotide of SEQ ID NO: 1, wherein the polynucleotide encodes
        a polypeptide with Rad50 activity.

Serial No. 09/538,396
Group Art Unit: 1638

14.   A polynucleotide which selectively hybridizes, under stringent hybridization conditions and a wash in 0.1X SSC at 60°C, to a polynucleotide of SEQ ID NO: 1.

15.   A polynucleotide comprising at least 30 contiguous nucleotides from a polynucleotide of claim 12.

⧖ NCBI

≡Protein

| PubMed | Nucleotide | Protein | Genome | Structure | PopSet | Taxonomy | OMIM | Books |

Search [Protein ▼] for [   ] [Go] [Clear]

Limits   Preview/Index   History   Clipboard   Details

[Display] [default ▼] [Save] [Text]   [Add to Clipboard]

☐ 1: AAD15407. putative RAD50 DN...[gi:4263721] Nucleotide, Related Sequences, PubMed, Taxonomy, BLink, LinkOut

```
LOCUS       AAD15407                1292 aa            linear   PLN 05-APR-2000
DEFINITION  putative RAD50 DNA repair protein [Arabidopsis thaliana].
ACCESSION   AAD15407
PID         g4263721
VERSION     AAD15407.1  GI:4263721
DBSOURCE    locus AC006223 accession AC006223.3
KEYWORDS    .
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (residues 1 to 1292)
  AUTHORS   Lin,X., Kaul,S., Rounsley,S.D., Shea,T.P., Benito,M.-I., Town,C.D.,
            Fujii,C.Y., Mason,T.M., Bowman,C.L., Barnstead,M.E.,
            Feldblyum,T.V., Buell,C.R., Ketchum,K.A., Lee,J.J., Ronning,C.M.,
            Koo,H., Moffat,K.S., Cronin,L.A., Shen,M., VanAken,S.E., Umayam,L.,
            Tallon,L.J., Gill,J.E., Adams,M.D., Carrera,A.J., Creasy,T.H.,
            Goodman,H.M., Somerville,C.R., Copenhaver,G.P., Preuss,D.,
            Nierman,W.C., White,O., Eisen,J.A., Salzberg,S.L., Fraser,C.M. and
            Venter,J.C.
  TITLE     Sequence and analysis of chromosome 2 of the plant Arabidopsis
            thaliana
  JOURNAL   Nature 402 (6763), 761-768 (1999)
  MEDLINE   20083487
   PUBMED   10617197
REFERENCE   2  (residues 1 to 1292)
  AUTHORS   Lin,X.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-MAR-2000) The Institute for Genomic Research, 9712
            Medical Center Dr., Rockville, MD 20850, USA
COMMENT     Method: conceptual translation.
FEATURES             Location/Qualifiers
     source          1..1292
                     /organism="Arabidopsis thaliana"
                     /cultivar="Columbia"
                     /db_xref="taxon:3702"
                     /chromosome="II"
     Protein         1..1292
                     /product="putative RAD50 DNA repair protein"
     CDS             1..1292
                     /gene="At2g31970"
                     /coded_by="join(AC006223.3:17700..17822,
                     AC006223.3:17919..18005,AC006223.3:18137..18304,
                     AC006223.3:18621..18796,AC006223.3:18911..19025,
                     AC006223.3:19249..19477,AC006223.3:19659..19837,
                     AC006223.3:19925..20095,AC006223.3:20187..20291,
                     AC006223.3:20381..20488,AC006223.3:20588..20740,
                     AC006223.3:20875..20978,AC006223.3:21144..21271,
                     AC006223.3:21551..21665,AC006223.3:21785..21914,
```

```
        AC0062     :22023..22232,AC006223.3:2242●22637,
        AC006223.3:22749..22950,AC006223.3:23535..23704,
        AC006223.3:23764..24100,AC006223.3:24481..24606,
        AC006223.3:24692..24744,AC006223.3:24855..25005,
        AC006223.3:25081..25143,AC006223.3:25254..25387,
        AC006223.3:25590..25722)"
ORIGIN
        1 mstvdkmlik girsfdpenk nvvtffrplt livgangagk ttiieclkvs ctgelppnar
       61 sghsfihdpk vagetetkaq iklrfktaag kdvvcirsfq ltqkaskmey kaiesvlqti
      121 nphtgekvcl syrcadmdre ipalmgvska ilenvifvhq desnwplqdp stlkkkfddi
      181 fsatrytkal evikklhkdq aqeiktfklk lenlqtlkda ayklresiaq dqertesskv
      241 qmleletsvq kvdaevhnke mmlkdlrklq dqvsiktaer stlfkeqqrq yaalpeened
      301 tieelkewks kfeerlallq tkirkmerem vdtettissl hnaktnymle isklqteaea
      361 hmllknerds tiqniffhyn lgnvpstpfs tevvlnltnr iksrlgelem dlldkkksne
      421 talstawdcy mdandrwksi eaqkrakdei kmgiskriee keierdsfef eistvdvkqt
      481 derekqvqve lerktkqnse rgfeskieqk qheiyslehk iktlnrerdv magdacdrll
      541 trideckdri rgvlkgrlpp ekdmkreivq alrsiereyd dlslksreae kevnmlqmki
      601 qevnnslfkh nkdtesrkry iesklqalkq esvtidaypk llesakdkrd drkrcynman
      661 gmrqmfepfe krarqehscp ccersftade easfikkqrv kasstgehlk alavessnad
      721 svfqqldklr avfeeysklt teiiplaekt lqehteelgq ksealddvlg isaqikadkd
      781 siealvqple nadrifqciv syqkqiedle ykldfrglgv ktmeeiqsel sslqsskdkl
      841 hgeleklrdd qiymerdisc lqarwhavre ekakaanllr dvtkaeedle rlaeeksqld
      901 ldvkylteal gplskekeql lsdyndmkir rnqeyeelae kkrnyqqeve allkasykin
      961 dcftryhdlk kgerlddiqe kqrlsdsqlq scearknela gelnrnkdlm rnqdqlrcni
     1021 ednlnyrttk akveeltrei esleeqilni ggiaaveaei vkilrererl lselnrcrgt
     1081 vsvyessisk nrvclkqaqy kdidkrhfdq liqlkttema nkdldryyna ldkalmrtht
     1141 mkmeeinkii relwqqtyrg qdmdyirihs dsegagtrsy sykvlmqtgd telemrgrcs
     1201 agqkvlasli irlalaetfc lncgilalde pttnldgpns eslagallri medrkgqcnf
     1261 qlivithder faqmiqqrqh aekyyrvakd dm
//
```

Revised: October 24, 2001.

NCBI   Nucleotide

PubMed   Nucleotide   Protein   Genome   Structure   PopSet   Taxonomy   OMIM   Books

Search Nucleotide ▼ for [   ]   Go Clear

Limits   Preview/Index   History   Clipboard   Details

Display default ▼ Save Text   Add to Clipboard

☐ **1: AF168748. Arabidopsis thali...[gi:7110147]**   Related Sequences, Protein, PubMed, Taxonomy, LinkOut

```
LOCUS       AF168748               4305 bp    mRNA    linear   PLN 04-MAY-2001
DEFINITION  Arabidopsis thaliana DNA repair-recombination protein (RAD50) mRNA,
            complete cds.
ACCESSION   AF168748
VERSION     AF168748.1  GI:7110147
KEYWORDS    .
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 4305)
  AUTHORS   Gallego,M.E., Jeanneau,M., Granier,F., Bouchez,D., Bechtold,N. and
            White,C.I.
  TITLE     Disruption of the Arabidopsis RAD50 gene leads to plant sterility
            and MMS sensitivity
  JOURNAL   Plant J. 25 (1), 31-41 (2001)
  MEDLINE   21097002
   PUBMED   11169180
REFERENCE   2  (bases 1 to 4305)
  AUTHORS   Gallego,M.E., Nagpal,P., Quatrano,R. and White,C.I.
  TITLE     The RAD50 homolog of Arabidopsis
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 4305)
  AUTHORS   Gallego,M.E., Nagpal,P., Quatrano,R. and White,C.I.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-JUL-1999) UMR 6547 - CNRS, Universite Blaise Pascal,
            24, Ave. des Landais, Aubiere 63170, France
FEATURES             Location/Qualifiers
     source          1..4305
                     /organism="Arabidopsis thaliana"
                     /cultivar="Columbia"
                     /db_xref="taxon:3702"
                     /chromosome="II"
                     /map="near TEn5"
                     /note="cloned from cell suspension culture"
     gene            1..4305
                     /gene="RAD50"
     CDS             146..4096
                     /gene="RAD50"
                     /note="similar to yeast RAD50"
                     /codon_start=1
                     /product="DNA repair-recombination protein"
                     /protein_id="AAF36810.1"
                     /db_xref="GI:7110148"
                     /translation="MSTVDKMLIKGIRSFDPENKNVVTFFRPLTLIVGANGAGKTTII
                     ECLKVSCTGELPPNARSGHSFIHDPKVAGETETKAQTKLREKTAAGKDVVCIRSFQLT
                     QKASKMEYKATESVLQTTNPHTGEKVCLSYRCADMDREIPALMGVSKAILENVIFVHQ
                     DESNWPLQDPSTLKKKFDDIFSATRYTKALEVIKKLHKDQAQEIKTFKLKLENLQTLK
                     DAAYKLRESIAQDQERTESSKVQMLELETSVQKVDAEVHNKEMMLKDLRKLQDQVSIK
```

```
            TAERST●  QQRQYAALPEENEDTIEELKEWKSKFEER  .GTKIRKMEREMVDTE
            TTISSLHNAKTNYMLEISKLQTEAEAHMLLKNERDSTIQNIFFHYNLGNVPSTPFSTE
            VVLNLTNRIKSRLGELEMDLLDKKKSNETALSTAWDCYMDANDRWKSIEAQKRAKDEI
            KMGISKRIEEKEIERDSFEFEISTVDVKQTDEREKQVQVELERKTKQNSERGFESKIE
            QKQHEIYSLEHKIKTLNRERDVMAGDAEDRVKLSLKKTEQENLKKKHKKIIDECKDRI
            RGVLKGRLPPEKDMKREIVQALRSIEREYDDLSLKSREAEKEVNMLQMKIQEVNNSLF
            KHNKDTESRKRYIESKLQALKQESVTIDAYPKLLESAKDKRDDRKREYNMANGMRQMF
            EPFEKRARQEIISCPCCERSFTADEEASFIKKQRVKASSTGEIILKALAVESSNADSVFQ
            QLDKLRAVFEEYSKLTTEIIPLAEKTLQEHTEELGQKSEALDDVLGISAQIKADKDSI
            EALVQPLENADRIFQEIVSYQKQIEDLEYKLDFRGLGVKTMEEIQSELSSLQSSKDKL
            HGELEKLRDDQIYMERDISCLQARWHAVREEKAKAANLLRDVTKAEEDLERLAEEKSQ
            LDLDVKYLTEALGPLSKEKEQLLSDYNDMKIRRNQEYEELAEKKRNYQQEVEALLKAS
            YKINEYHDLKKGERLDDIQEKQRLSDSQLQSCEARKNELAGELNRNKDLMRNQDQLRR
            NIEDNLNYRTTKAKVEELTREIESLEEQILNIGGIAAVEAEIVKILRERERLLSELNR
            CRGTVSVYESSISKNRVELKQAQYKDIDKRHFDQLIQLKTTEMANKDLDRYYNALDKA
            LMRFHTMKMEEINKIIRELWQQTYRGQDMDYIRIHSDSEGAGTRSYSYKVLMQTGDTE
            LEMRGRCSAGQKVLASLIIRLALAETFCLNCGILALDEPTTNLDGPNSESLAGALLRI
            MEDRKGQENFQLIVITHDERFAQMIGQRQHAEKYYRVAKDDMQHSIIEAQEIFD"
BASE COUNT      1544 a     733 c     999 g    1029 t
ORIGIN
       1 gttccttctt cttccattgt cgagttttca tgattttccc gcaaaaccct tctcgtttat
      61 gggtttaacc ttcctgtgtc accaaaattt tctccgtaac ttatcacctc cactttttca
     121 tatctgtaat cgctgaaggt aaaagatgag tacggtcgat aaaatgttga tcaagggtat
     181 aagaagcttc gacccagaaa ataagaacgt tgttactttc ttcagacctt taaccctaat
     241 cgtcggcgct aacggagccg gaaaaactac gattatagag tgtttaaagg tgtcttgtac
     301 cggagagttg ccaccgaacg cgaggtctgg tcatagcttc attcatgacc ctaaggttgc
     361 tggagaaacg gagacaaaag cacaaataaa gctgagattc aagacagctg cagggaagga
     421 tgtggtatgt ataaggtcat ttcagttgac gcaaaaggct tcaaagatgg agtacaaggc
     481 tattgagagt gtgcttcaaa ccattaatcc tcacactgga gagaaagtat gtctcagcta
     541 cagatgcgct gacatggata gggagatacc agctttgatg ggtgtgtcaa aagctatttt
     601 agaaaacgtc atatttgtgc atcaggatga gtctaactgg ccactgcaag atccttctac
     661 tttgaaaaag aagtttgatg atatttttctc agccaccaga tatacaaagg ctctggaagt
     721 tattaagaaa cttcacaagg atcaagctca ggagataaaa actttcaagt taaaattgga
     781 aaatctgcag accttgaaag acgcagctta taagcttcgt gagagcattg ctcaagatca
     841 agaaaggaca gaatcttcaa aagttcagat gttggaactg gaaaccagtg ttcagaaagt
     901 tgatgctgaa gttcataata aggaaatgat gttaaaagat ctgagaaagc tccaggacca
     961 agtatcaata aaaactgcgg aaaggagcac cttgtttaag gagcagcaga gacagtatgc
    1021 agcattacct gaggaaaatg aagatactat tgaagaattg aaagaatgga aaagcaagtt
    1081 tgaagaaaga cttgctcttt tggggaccaa aattcgaaaa atggagagag agatggttga
    1141 cacagaaact acaatttctt cgcttcacaa tgcaaaaact aactatatgc tggagatcag
    1201 caagctacaa acagaagctg aagctcatat gctgctgaag aatgaaagag attctaccat
    1261 ccagaacatt tttttccatt acaatttagg aaatgttcct agtaccccct tcagtactga
    1321 ggtagttttg aatctcacaa atcgaataaa atcaagactg ggtgaacttg agatggattt
    1381 gctcgataag aagaaatcaa atgaaactgc tctaagtacg gcttgggatt gctatatgga
    1441 tgctaatgac cgttggaaaa gcatcgaagc tcagaaacga gctaaagatg aaatcaagat
    1501 gggcatatca aaacgcatag aagagaaaga aattgagcgt gattcatttg aattcgaaat
    1561 ttccactgtg gatgttaaac aaactgatga agagagaaa caagtgcaag ttgagcttga
    1621 gagaaagaca aaacagaatt ctgaaagggg attgaatcc aaaatagaac agaaacagca
    1681 tgaaatatac agcttggagc ataaaatcaa gacactgaat cgagaaagag acgttatggc
    1741 tggtgatgct gaggatagag tgaaattgtc tttaaagaag acggaacagg agaatcttaa
    1801 gaagaaacac aaaaagataa tagatgagtg caaggatcga attagaggag tgctgaaggg
    1861 gagattacca cccgagaagg atatgaagag ggaaatagtt caagctttga ggtcaattga
    1921 aagggaatat gatgacctca gtttgaaatc ccgagaagca gaaaaggagg tcaacatgtt
    1981 gcagatgaag atacaagaag taaacaatag cctattcaag cataacaagg atacagaatc
    2041 aagaaagaga tacattgagt ccaagcttca ggccttaaaa caggaatcag ttaccattga
    2101 tgcttatccc aaattgttgg agtcggctaa ggacaaaaga gatgaccgta acgcgaata
    2161 taacatggca aatggtatgc gtcagatgtt tgaacctttt gagaaacgag cacgtcaaga
    2221 gcattcttgt ccttgctgcg agcgctcttt tacagctgat gaagaggcta gcttattaa
    2281 gaagcaaagg gtaaaggctt caagcacagg ggagcatctt aaagcattag cagtggagtc
    2341 ttccaatgct gactctgtat ttcaacagct agacaaacta cgggcagttt ttgaagaata
    2401 ctcgaaacta actactgaaa ttattcctct cgctgagaaa actttgcaag agcatacaga
    2461 agagctgggt cagaagtctg aagctcttga tgatgtatta ggaatctctg cacagataaa
    2521 agctgacaaa gactcgattg aagcactggt gcagcccctc gaaaatgctg acaggatctt
```

```
2581 tcaggaaatc gtatcct..c aaaagcaaat tgaagatctt gaatat..c ttgattttag
2641 gggtcttggg gttaaaacta tggaagaaat ccaatcagaa ttaagtagcc tccaaagcag
2701 caaggataaa ttgcacggtg agctggagaa acttagagat gaccagattt acatggaacg
2761 tgatatatca tgcctccagg ccagatggca tgcagtaagg gaggagaagg ccaaggcagc
2821 caatttgtta cgggacgtta caaaggcaga agaagacctg gagcgtttag ccgaggagaa
2881 aagtcaactt gacctcgatg tgaagtattt gaccgaggcc ctgggccctt tatccaagga
2941 aaaagagcaa ttactgagtg actacaatga tatgaagatt agacgtaatc aagagtatga
3001 agaactggct gagaaaaaga gaaactacca acaggaagtt gaggcattac tcaaggccag
3061 ttataaaata aatgagtatc atgatctgaa gaaaggggag cgcttagatg atattcaaga
3121 aaagcagcgg ttatctgatt ctcagcttca gagctgtgaa gctagaaaga atgaacttgc
3181 aggtgaactg aacagaaata aagacttgat gcggaatcaa gatcaactga gaagaaatat
3241 agaggataac ttgaactacc ggacaacaaa agcaaagtt gaggagctta cacgtgagat
3301 tgaatcattg gaagagcaaa tattgaatat tggtgggatt gccgcagttg aagctgaaat
3361 tgtaaagata ttgcgggaaa gggaaagact tctttcagag ctgaaccggt gtcgtggaac
3421 agtgtctgtt tacgagagca gtatttcaaa gaacagagtg gagctaaaac aggcacaata
3481 caaggacata gataagcggc actttgatca actgatccag ctaaagacaa ctgaaatggc
3541 aaataaagac ttggatagat actacaatgc ccttgacaaa gcactaatgc gcttccacac
3601 aatgaaaatg gaggaaataa ataagattat aagggaactg tggcagcaga catacagagg
3661 tcaagatatg gattacatac gaatacactc agattccgaa ggtgcaggca cccgctctta
3721 cagctacaag gttcttatgc agactggtga tacagaactt gaaatgagag gaagatgcag
3781 tgcaggtcaa aaggttcttg cttcattgat aattaggttg gctttagccg agacattttg
3841 cttaaactgt ggaatattag cccttgatga gccaactaca aacctggatg ggcccaattc
3901 agaatctctt gcaggagctc ttcttaggat catggaagac agaaagggcc aagagaattt
3961 ccaactcata gtcattaccc acgatgaacg ctttgctcaa atgattggcc aaaggcaaca
4021 tgctgaaaaa tattaccggg tcgcaaaaga tgatatgcaa cacagcataa tcgaggccca
4081 agagatcttt gattgataaa ccagtggctt gttctataaa ccaacataac agttcagcta
4141 atgtttaacg gtaagctcat ctttttgaaa agctacgtgt agctttcttg gctctgttat
4201 atacagctct acacattgtt tgattagatg tgtgcaaaaa tttctgttga attgaaaaat
4261 cagtatgtag aggagacaaa aaaaaaaaaa aaaaaaaaaa aaaaa
//
```

Revised: October 24, 2001.

NCBI Sequence Viewer

Search Nucleotide ▼ for _____   Go Clear

Limits   Preview/Index   History   Clipboard   Details

Display default ▼ Save Text   Add to Clipboard

☐ 1: X14814. Yeast RAD50 gene ...[gi:4272]   Related Sequences, Protein, PubMed, Taxonomy

```
LOCUS       SCRAD50              4775 bp    DNA        linear    PLN 12-SEP-1993
DEFINITION  Yeast RAD50 gene for 153 kD protein.
ACCESSION   X14814
VERSION     X14814.1  GI:4272
KEYWORDS    DNA repair; DNA-binding protein; meioticrecombination; RAD50 gene.
SOURCE      baker's yeast.
  ORGANISM  Saccharomyces cerevisiae
            Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
            Saccharomycetales; Saccharomycetaceae; Saccharomyces.
REFERENCE   1  (bases 1 to 4775)
  AUTHORS   Alani,E.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-MAR-1989) Alani E., Harvard University, 7 Divinity
            Avenue, Cambridge MA 02138, U S A
REFERENCE   2  (bases 1 to 4775)
  AUTHORS   Alani,E., Subbiah,S. and Kleckner,N.
  TITLE     The yeast RAD50 gene encodes a predicted 153Kd protein containing a
            purine nucleotide binding domain and two large heptad repeat
            regions
  JOURNAL   Genetics 112, 47-57 (1989)
REFERENCE   3  (bases 1 to 4775)
  AUTHORS   Raymond,W.E. and Kleckner,N.
  TITLE     Expression of the Saccharomyces cerevisiae RAD50 gene during
            meiosis: steady-state transcript levels rise and fall while
            steady-state protein levels remain constant
  JOURNAL   Molecular & general genetics : MGG. 238 (3), 390-400 (1993)
  MEDLINE   93261422
   PUBMED   8492807
COMMENT     Data kindly reviewed (25-SEP-1989) by Alani E.
FEATURES             Location/Qualifiers
     source          1..4775
                     /organism="Saccharomyces cerevisiae"
                     /strain="RE821"
                     /db_xref="taxon:4932"
                     /chromosome="chromosome 14"
                     /clone="pSG205"
                     /clone_lib="ARSCEN"
     misc_feature    532..541
                     /note="region of transcription start"
     CDS             558..4496
                     /note="153 kD protein (AA 1 - 1312)"
                     /codon_start=1
                     /protein_id="CAA32919.1"
                     /db_xref="GI:4273"
                     /db_xref="SWISS-PROT:P12753"
                     /translation="MSAIYKLSIQGIRSFDSNDRETIEFGKPLTLIVGMNGSGKTTII
                     ECLKYATTGDLPPNSKGGVFIHDPKITGEKDIRAQVKLAFTSANGLNMIVTRNIQLLM
                     KKTTTTFKTLEGQLVAINNSGDRSTLSTRSLELDAQVPLYLGVPKAILEYVIFCHQED
                     SLWPLSEPSNLKKKFDEIFQAMKFTKALDNLKSTKKDMSVDIKLLKQSVEHLKLDKDK
                     SKAMKLNIIQLQTKIDQYNEEVSEIESQLNEITEKSDKLFKSNQDFQKILSKVENLKN
```

```
                    TKLSISD**KRLSNSIDILDLSKPDLQNLLANFSKVLMDK**QLRDLETDISSLKDRQ
                    SSLQSLSNSLIRRQGELEAGKETYEKNRNHLSSLKEAFQHKFQGLSNIENSDMAQVNH
                    EMSQFKAFISQDLTDTIDQFAKDIQLKETNLSDLTKSITVDSQNLEYNKKDRSKLIHD
                    SEELAEKLKSFKSLSTQDSLNHELENLKTYKEKLQSWESENIIPKLNQKIEEKNNEMI
                    ILENQIEKFQDRIMKTNQQADLYAKLGLIKKSINTKLDELQKITEKLQNDSRIRQVFF
                    LTQEFQRADLEMDFQKLFINMQKNIAINNKKMHELDRRYTNALYNLNTTEKDLQDNQK
                    SKEKVIQLLSENLPEDCTIDEYNDVLEETELSYKTALENLKMHQTTLEFNRKALEIAE
                    RDSCCYLCSRKFENESFKSKLLQELKTKTDANFFKTLKDTVQNEKEYLHSLRLLEKHI
                    ITLNSINEKIDNSQKCLEKAKEETKTSKSKLDELEVDSTKLKDEKELAESEIRPLIEK
                    FTYLEKELKDLENSSKTISEELSIYNTSEDGIQTVDELRDQQRKMNDSLRELRKTISD
                    LQMEKDEKVRENSRMINLIKEKELTVSEIESSLTQKQNIDDSIRSKRENINDIDSRVK
                    ELEARIISLKNKKDEAQSVLDKVKNERDIQVRNKQKTVADINRLIDRFQTIYNEVVDF
                    EAKGFDELQTTIKELELNKAQMLELKEQLDLKSNEVNEEKRKLADSNNEEKNLKQNLE
                    LIELKSQLQHIESEISRLDVQNAEAERDKYQEESLRLRTRFEKLSSENAGKLGEMKQL
                    QNQIDSLTHQLRTDYKDIEKNYHKEWVELQTRSFVTDDIDVYSKALDSAIMKYHGLKM
                    QDINRIIDELWKRTYSGTDIDTIKIRSDEVSSTVKGKSYNYRVVMYKQDVELDMRGRC
                    SAGQKVLASIIIRLALSETFGANCGVIALDEPTTNLDEENIESLAKSLHNIINMRRHQ
                    KNFQLIVITHDEKFLGHMNAAAFTDHFFKVKRDDRQKSQIEWVDINRVTY"
     misc feature  642..707
                    /note="nucleotide-binding domain"
     misc feature  1089..1820
                    /note="heptad repeat"
     misc feature  2784..3542
                    /note="heptad repeat"
BASE COUNT        1827 a      768 c      895 g     1285 t
ORIGIN
        1 ggatatctaa caagccatca aaagcataat gttttttgagt actcgataga attatgtatt
       61 gcggacttta gtagaatagg aaacaatgcg ctatacgtgc gcagtagctc atttcgacca
      121 taaactaatt gtagattttt tgcgtttctg gtgctattca ttgaaacaag aaaccgatag
      181 tacttccact tacaatacta atacttgcgg taatgttaca atccttcact tattattttc
      241 ggagaataat aaccatgcat cttgcaatac tttcacggct ttgcttgtct cgtgcgctag
      301 taactgcagt agaacccacc catatgacaa tgcaaatatt aatctaactt agtcaaaaaa
      361 aaaaaatgcc ttcacctcgt ttgtcttcat ctgagataag gttccggttt tcatcttgtc
      421 cctgtttgtt gcctttgttc ttgcctcatc gcttctaaag cggctttcaa gctttgatct
      481 ttaaatggaa atctttaatt ctaagaagtg gtagcaacca ttgagaggca aaaacaaggg
      541 aacgacggaa agcaggcatg agcgctatct ataaattatc tattcagggc atacggtctt
      601 ttgactccaa tgatagggaa actattgaat ttggcaagcc tctgacttta atagtcggca
      661 tgaatggttc aggtaaaacg actatcatcg aatgtttaaa gtacgctacc acaggtgatc
      721 taccgcccaa cagcaaggga ggagtattca ttcatgaccc gaagataact ggtgaaaagg
      781 acattagagc tcaggtcaaa ctggcgttta cgagtgccaa tggactcaat atgattgtca
      841 ccagaaatat tcagttgcta atgaaaaaga ctactactac atttaagact ttagaaggcc
      901 agttggtcgc tataaataat agtggtgacc gcagtacttt gtccacccgg tctctggaat
      961 tagacgcaca agtacccctg tatctgggag ttccaaaagc gatcttggaa tacgtgatat
     1021 tctgtcacca agaagacagc ctgtggccac tgagtgaacc atccaatctt aaaaagaagt
     1081 ttgatgaaat ttttccaagcg atgaaattta caaaggctct agacaatttta aaatccataa
     1141 agaaagacat gtccgtggat atcaagttat taaagcaatc agtagaacac ttaaagttgg
     1201 ataaagacag atcgaaagca atgaagctca atatccacca attgcaaact aagatagatc
     1261 agtataatga agaagtgtct gaaattgaat ctcaattgaa tgaaatcact gagaaatccg
     1321 ataaactgtt taaatctaat caagatttcc aaaaaaatact atccaaagta gagaacttga
     1381 aaaatacaaa actctcaatt agtgatcaag tgaaaagatt gtcaaactca attgatattc
     1441 tcgatttatc gaagccagat ttgcaaaact tattggcaaa cttctctaaa gtcttgatgg
     1501 ataagaataa tcagttgaga gatttggaaa ccgatatttc aagtttgaaa gaccgtcagt
     1561 catccttgca gagcctttct aattccctca tcaggcggca gggtgagtta gaagctggta
     1621 aggagacata cgaaaagaac agaaatcacc tttcatcgct aaaagaagcc ttccaacata
     1681 agttccaagg ccttcaaat attgaaaaca gtgatatggc tcaagtaaat cacgaaatgt
     1741 cacagttcaa ggctttcatc tctcaggatt taactgacac tatagatcaa tttgcaaagg
     1801 atattcaact aaaggaaaca aatttatccg acttgataaa atcaataacc gtcgattcac
     1861 agaatttgga atataataaa aaggacagat ctaaattaat ccatgactct gaagagcttg
     1921 ctgagaaatt gaaatcattt aaaagttggt ctactcaaga tagcctcaat catgaactgg
     1981 aaaacttaaa aacatacaaa gaaaagctgc aatcttggga aagtgagaat attataccaa
     2041 aattgaacca gaaaatagaa gagaaaaata atgagatgat tatcttagag aatcaaattg
     2101 agaaatttca agatcgcata atgaaaacga atcaacaagc ggacttatac gccaaactag
     2161 gactaataaa gaagtccatt aatactaaat tagacgagtt acaaaagatt acggaaaaat
```

```
2221 tacagaatga ttctag    a aggcaagtat ttcccctcac ccagga   t caaagagctg
2281 atttagaaat ggattttcag aaattattca tcaatatgca aaagaatata gcgatcaaca
2341 acaaaaagat gcacgaattg gatagaagat atacaaacgc tttgtacaat ttgaacacca
2401 ttgaaaagga tttacaggat aatcagaagt cgaaagaaaa agtgatacaa ctactaagcg
2461 aaaacttacc agaggattgt actattgatg aatataatga tgtcctggaa gaaacggagc
2521 tctcctataa gactgcgctg gaaaacttga aaatgcatca aactacttta gaatttaata
2581 gaaaagcatt agagatcgct gaacgtgaca gctgctgtta tttgtgttct agaaagtttg
2641 aaaatgaatc gtttaaaagt aagttgttac aagaactaaa aacaaaaaca gacgcaaatt
2701 ttgaaaaaac tttgaaagac actgtccaaa acgaaaagga atatcttcat agtttaaggc
2761 tactggagaa acatattatt acactaaatt ctataaatga aaagattgat aattcacaga
2821 agtgccttga aaaagctaag gaggaaacaa aaacttccaa atcaaaacta gatgaattgg
2881 aagtagattc cacaaaattg aaagatgaga aagagttagc tgagtctgaa attcgtcctt
2941 tgattgagaa gtttacttac ttagaaaagg aacttaagga tcttgaaaac agttctaaaa
3001 cgatatctga ggagctatcg atttataata ccagtgaaga tggtattcaa actgttgatg
3061 aattaaggga ccaacaaagg aagatgaatg attctttacg tgaactgagg aaaaccattt
3121 cagatttgca aatggaaaaa gatgaaaagg tgagagaaaa ttcaagaatg attaacttga
3181 ttaaagaaaa ggaattaaca gtttctgaaa ttgaatcatc tttaacacag aaacaaaata
3241 ttgatgattc tataaggtca aaaagggaaa acatcaacga tatcgattct agagtaaagg
3301 agttagaagc acgtattatt tcattaaaaa acaaaaagga tgaagcgcaa agtgttctag
3361 acaaagtaaa aaatgaacgt gatattcaag tacgcaacaa acaaaaaacc gttgcagata
3421 ttaatcggtt aatagataga tttcagacca tttataacga agtggttgat tttgaagcta
3481 agggcttcga tgaattgcaa acaacaataa aagaactaga attgaacaaa gctcaaatgc
3541 tggaactaaa ggagcagctt gatctcaagt caaatgaagt taacgaagaa aagagaaaac
3601 ttgctgactc aaataatgag gagaaaaact tgaagcagaa cttagaacta attgaattga
3661 aatcacaact tcagcacatt gaatctgaga taagtagatt agacgttcaa aatgcagaag
3721 ctgaaaggga caaatatcaa gaggaatcac taagactaag aacaagattt gaaaagctga
3781 gctctgaaaa cgcaggtaaa ttaggtgaaa tgaaacaatt acagaatcaa atagattcat
3841 tgactcatca actgcggact gattacaaag atatcgagaa aaattaccat aaagaatggg
3901 ttgaattaca gacaagatcc tttgttactg acgacataga tgtctattca aaagcattag
3961 atagtgctat tatgaaatat catggattga aaatgcaaga tataaacagg ataattgatg
4021 aattatggaa gcgtacgtat agtggtactg atatagacac catcaaaatt agatcagatg
4081 aagttagtag tacagttaag ggaaaatcat acaactatcg tgttgtaatg tacaagcagg
4141 atgttgaatt ggatatgaga ggacgatgtt ccgcgggtca aaaagttctt gcatcgatta
4201 ttatcaggct agctctatct gaaacatttg gcgcaaattg tggcgtaatt gcactagacg
4261 aacctaccac caatttagat gaagaaaaca tcgaaagttt agcaaaatct ttacacaata
4321 taattaatat gagaagacat caaaagaatt tccagctaat cgtgatcact catgatgaga
4381 aatttttggg tcatatgaat gcagcggcat ttactgatca cttcttcaaa gtcaagagag
4441 atgataggca aaaatcacaa attgagtggg tcgatattaa cagagtcact tattgaaaag
4501 accccataat atctacgaag ggatagactt tgattgatta attagaggtg ctgaaatact
4561 tcgtcgacga gtttgtgata acgcttctgt tgcttagcca tatctaagtg tatatgatat
4621 ttaacgcaaa cgaattttct tagtaatatg ctatcgtttc caagtcggta ttcagcaaga
4681 gattcgaatt tttcgctatc attactaata tctctctctg ttataacgtc agctagttgt
4741 ttcttttcct ttgcagactt agcaagcacg taagc
```

//

Revised: October 24, 2001.

NCBI
PubMed   Nucleotide   Protein   Genome   Structure   PopSet   Taxonomy   OMIM   Books

Search [Nucleotide ▼] for [   ]   [Go] [Clear]

Limits   Preview/Index   History   Clipboard   Details

[Display] [default ▼] [Save] [Text]   [Add to Clipboard]

☐ 1: Z75312. C.elegans mRNA fo...[gi:2687854]   Related Sequences, Protein, Taxonomy, LinkOut

```
LOCUS           CERAD50                   4121 bp    mRNA     linear   INV 11-DEC-1997
DEFINITION      C.elegans mRNA for RAD50.
ACCESSION       Z75312
VERSION         Z75312.1  GI:2687854
KEYWORDS        RAD50.
SOURCE          Caenorhabditis elegans.
  ORGANISM      Caenorhabditis elegans
                Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
                Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
REFERENCE       1  (bases 1 to 4121)
  AUTHORS       Offenberg,H.H. and Heyting,C.
  JOURNAL       Unpublished
REFERENCE       2  (bases 1 to 4121)
  AUTHORS       Offenberg,H.H.
  TITLE         Direct Submission
  JOURNAL       Submitted (10-JUL-1996) Offenberg H.H., Agricultural University,
                Genetics, Dreyenlaan 2, Wageningen, 6703 HA The Netherlands
FEATURES             Location/Qualifiers
     source          1..4121
                     /organism="Caenorhabditis elegans"
                     /strain="CB1489 him-8(e1489)"
                     /db_xref="taxon:6239"
                     /sex="male/hermaphrodite"
                     /tissue_type="whole animal"
                     /clone_lib="Yuji Kohara cDNA (unpublished)"
                     /dev_stage="varied"
     gene            31..3927
                     /gene="RAD50"
     CDS             31..3927
                     /gene="RAD50"
                     /function="DNA repair and recombindation protein"
                     /codon_start=1
                     /product="RAD50 homologue ceRAD50"
                     /protein_id="CAA99730.1"
                     /db_xref="GI:2687855"
                     /db_xref="SPTREMBL:O44199"
                     /translation="MAKFLRLHIRGIRSVGDEDHDVHKIDFLSPCTLISGPNGTGKTT
                     TIEALNFVTTGQMPTQKKQNFIHSTDVARKTRVDASVTLEFIDVKGRECTAVRRLVVT
                     SGTKAAALAEEITLAIKYPDGTVNTLSSKVCDFNTALLKHLGVPRAVFKYVIFCHQED
                     STWPLSEPKELKKRFDDIFQLTKFVKAQERMKKIVLDFKKEMQTHEMSKQLYETHVRD
                     KLVARQNQEECERKISKRKEETDELKERKANGQKKIEEMRTSIHELEDTLTSFKKTEL
                     ERQNLKKQLSLIRVEPYFGTEEELKREIEEFRGSEGRSYGEERARIQKKIGKNNQERQ
                     ELSQKKTEFENRISSLKAEVIHCQSLKYDLERLENQLRSELDLEHDADIDIEIDNAIT
                     LKIRCMSDKARMIAKNCAELQSNLRTAQEAATKTEVFMKTLQNEKVKLEKEVEQLKFK
                     IKQGQNATAGMKDLLKKEEALRKSLADLPLLDENALTECKLKREKYLKQLDILKKKCA
                     EAEKNAEKDRRKESLKQTLSIARKKMTAYQRIYDNNWQGLLGQAPDFPWTPILSKTEII
                     KLRNDKKIMEEDLRDVQLNVQKLETMQHQYRKQEESLTAQELKLSENIFEACSCEAEE
                     VSEKLENLRKRLKKARKDLAPLSAKSNLYDSYIEESKSSGCCPLCDRDFKTKKEINEF
                     SKKLENMTLSFPTEQEELEKLVSKLEKEEIIIVKAEGQANELQRIVKELKEVRFKNRK
                     LSTEMAEEKSNLSKNFKQLETVNAKLKLAEDLQTDVGVIQQLYEQTEENEKRYEQLVS
```

ESDSSDG●●YTELRKKVEDKDEEYRKIVQEGEELQKCSEE●●KLQSKLNELGTHRVSL
GEAAAQAGAFAEQLETKIKEIQECTTAISQKRNEDLPDAQFKKDDLTRNVSSKEEEKK
KAEMEVQMMKKELDQKIFHRKSLFKKVQEGGLCERQLMDKENNIATLNASLEENQQRQ
KRFEEDLRSFDSSHQRESILKDQLTRMIIENKIKELKRTLATFDGQINEDRITEQKQA
YNKLQNELRLIGNEEVKIYTQMQEYEKQKKIAEAKLSTKECQNAESNYRDAIIELAIT
KESISDLTKYRNCLDASLTQFHSEKMGRVNGIIDDLWRKVYNSTDITTIRIRSDATSE
TSSKKVAYEYNVMMVHETGTEVEMRGRCSAGQKMLASLLIRIALAEVFGGSCSMIALD
EPTTNLDESKVEGMAIVLADTIAERRGFDENGKLRGRDMQMVVITHDERLVNRITISC
RPEYIYCLGKDEHGTSFLSKRYPDGTVKRVNTKRRF"

BASE COUNT   1573 a   702 c   889 g   957 t
ORIGIN
        1 tatttgctta tcgaagcctg ggaattgctt atggcgaaat ttttacgcct acacatcaga
       61 ggaattcgaa gtgtcggtga tgaggaccac gatgttcaca aaattgattt cctctctccg
      121 tgcacactaa tcagtggacc aaatggaaca ggaaaaacga ccaccatcga ggcactcaac
      181 tttgtaacca ccggacaaat gcctacacag aagaaacaga actttattca cagcacagat
      241 gtagcacgca aaactcgtgt tgatgcctca gttactctcg aattcattga tgtcaaagga
      301 cgagagtgta cagctgtacg gagacttgtt gtgacatcag gaacgaaagc tgcagcacta
      361 gcagaagaac acacattagc tataaaatat cctgatggca cagtcaatac tctgagcagc
      421 aaagtttgcg acttcaacac tgcactttta aaacatttag gcgtcccgag agcagttttc
      481 aaatatgtaa ttttctgtca tcaagaagat tcgacttggc cactcagtga accaaaagaa
      541 ctgaaaaaac gttttgatga catcttccaa ctgacaaaat ttgtgaaagc acaggaacga
      601 atgaaaaaaa ttgtgctgga ttttaagaaa gaaatgcaaa cccacgagat gtctaaacaa
      661 ctctacgaga ctcatgttag ggataaactg gtcgctagac aaaaccaaga agaatgtgag
      721 agaaaaatat cgaaacggaa agaggaaact gatgagctga agaacgaaa agccaatgga
      781 cagaagaaga tagaagagat gagaacgtct attcacgaac tcgaggacac attaacttcg
      841 ttcaagaaaa ctgaattgga gaggcaaaac ttgaagaaac agctttctct cattcgtgtg
      901 gagccctatt ttggaactga agaagagttg aagagagaaa ttgaagaatt ccgtggttct
      961 gaaggacgca gctatggtga agaacgggcc agaatccaga aaaaaattgg taaaaataat
     1021 caggaacgac aggagttgtc gcaaaagaaa actgaattcg agaacagaat ttcaagttta
     1081 aaagctgaag ttattcactg ccaatcactg aagtatgacc tggaacgact agaaaatcaa
     1141 cttcgttccg aactcgatct tgaacacgat gcggatattg acattgaaat tgataacgct
     1201 atcacattga aaattcgagg aatgtctgat aaagcacgaa tgattgcaaa gaactgtgcc
     1261 gaacttcaat ccaatctgcg aacagcacaa gaagcagcta caaaaattga agtagaaatg
     1321 aaaactcttc aaaacgaaaa agtcaaactg gagaagaag tagaacagct gaaattcaaa
     1381 atcaaacaag gacagaatgc cactgcaggc atgaaggatc ttctgaaaaa agaagaagct
     1441 ctgagaaaaa gtcttgctga ccttccatta ctcgacgaga atgcactgac agaatgcaaa
     1501 ctcaagcgag agaaatacct aaaacaattg gacattctga aaagaagtg cgccgaggct
     1561 gaaaagaatg ctgagaaaga cagagaaaaa gaatccttga aacaaactct gtctattgct
     1621 cgaaaaaaga tgactgcata tcaaaggata tatgataaca attggcaagg attgattgga
     1681 caggctcccg attttccatg gactccaatt ctatcgaaga ctttccacaa attgaggaat
     1741 gataagaaaa tcatggaaga agatcttcga gatgttcaat aaacgtcca aaagcttgaa
     1801 accatgcaac atcagtacag gaagcaagag gaaagtctaa cagcacagga attgaagttg
     1861 agtgaaaaca ttttcgaagc ttgttcgtgt gaagctgaag aagtctccga gaaattggaa
     1921 aatcttcgaa agcgtctgaa gaaagcccga aaagatttag cgcctctcag tgcaaaatca
     1981 aatctctacg attcttatat cgaagaaagc aaaagcagtg gttgctgtcc actttgtgat
     2041 cgtgacttca agacaaagaa agaaataaat gaattctcaa agaagcttga aaatatgaca
     2101 cttagtttc ccactgaaca ggaagaattg gagaaattag tttcgaaatt ggaaaaagaa
     2161 gaaataatta ttgtgaaagc agaaggacaa gcaaacgaac tgcagagaat agtaaaagaa
     2221 ctaaaggaag ttcgagagaa aaacagaaag ctgtctactg aaatggcgga agagaaatcg
     2281 aatctttcta aaaacgaaaa acaactggaa accgtgaatg ctaaactcaa attagctgaa
     2341 gatttgcaaa cagatgttgg tgtaattcaa caactctatg aacaaacaga agagaacgaa
     2401 aagagatatg agcaacttgt ttctgaatct gattcttcgg atggcctcag ctacaccgaa
     2461 cttcggaaga aagttgaaga caaagatgaa gaatatcgta agattgttca agaaggagaa
     2521 gaacttcaga agtgttctga ggaaagaaat aaacttcaat caaaactgaa tgagctcgga
     2581 actcaccgag tttcttgg agaagctgct gcacaggctg gtgctttcgc agaacaattg
     2641 gaaactaaga taaagagat tcaagagtgt attacggcaa tctcgcaaaa aagaaatgaa
     2701 gatcttccgg atgctcagtt caagaaagat gatctcacaa gaaatgtttc cagtaaagaa
     2761 gaagagaaga aaaagctga aatggaagtg cagatgatga aaaaggagtt ggatcaaaaa
     2821 attttccatc ggaaatcgct gtttaaaaaa gtccaggaag gcggtctgtg cgaaagacaa
     2881 ctcatggata aggaaaataa tattgcaact ctgaatgcca gcttagaaga aaatcagcaa
     2941 cgccaaaaac gtttcgaaga agatctccgt tcgttcgata gttcacatca acgcgagtca
     3001 attctaaaag atcaacttac tcgaatgatt atcgaaaata agataaaaga gttgaaaaga
     3061 acattggcca cattcgatgg acaaataaat gaagatcgca taactgaaca aaaacaagct

```
3121 tataacaaac ttcaaa   a acttcgcttg attggaaacg aagaag   a aatatacact
3181 cagatgcaag agtacgagaa gcaaaaaaaa attgcggaag caaaattgtc aacgaaagag
3241 tgtcagaatg ctgaatcgaa ttatcgtgat gctatcatag aactggcaat tactaaagaa
3301 tctatctctg atttaacgaa gtatagaaat tgtcttgatg cttcactcat ccaattcat
3361 tctgaaaaaa tgggcagggt caatggaatc attgatgatc tatggagaaa agtatataat
3421 tccacggata ttactacaat caggattaga tcagatgcta cttctgagac ttcttcaaag
3481 aaagtcgcat acgaatacaa tgttatgatg gttcatgaaa ctggcaccga agttgagatg
3541 cgtggacgct gcagtgctgg tcaaaaaatg cttgcttcgt tgcttattcg cattgcactc
3601 gcagaagtat tcggtggttc atgctcgatg attgcgttgg atgaaccaac aacaaatttg
3661 gatgagagca aagtggaagg aatggctatc gttttggcag atataattgc tgaaagacgc
3721 ggattcgatg aaaatggaaa gttgagagga agagatatgc agatggtagt aatcactcac
3781 gatgagcgac tcgtcaacag gatcacaatc agttgtcgtc cagagtatat ttattgtttg
3841 gggaaggacg agcacggaat aagtttcctt tcgaaacgtt atccagatgg aacagtgaaa
3901 agagttaata cgaagagacg gttctagaat ggtattgaag aatctcaata gtacaaataa
3961 tattttacgg tctcaaaatt gattcgctca ccttttaatt atgttttatg ctttgtcaga
4021 tcattcccga gtttccaaat gttgttacct catgttttat attttacgag atgaacccga
4081 tgcatcggcc gatgtgaaat gaaaaaaaaa aaaaaaaaa a
```

//

Revised: October 24, 2001.

```
              1351                        1380
   U63139aa   R KKNIDQCS E VKCS SSL GFNVH-----
   Z75311aa   R KKNIDQCS E VKCS SSL GFNVH~~~~~
 NM022246aa   RVKKNIDQCS E VKSS NSL GSYVH~~~~~
   U66887aa   RVKKNMDQCS E VKCS SSL GSYVH~~~~~
 AAD15407aa      kodm~~~ ~~~~~~~~~ ~~~~~~~~~~
 AF168748aa   RV Kd MQHS  IF QEIFD~ ~~~~~~~~~~
   1116sid2   RV KDKNQHS  IE QEIFD~ ~~~~~~~~~~
   X14814aa    VK D RQKS QEEWV NRV TY~~~~~~~~
   Z75312aa   CLG DEHGIS F SKRYPDGT VKRVNTKRRF
```

NCBI

| PubMed | Nucleotide | Protein | Genome | Structure | PopSet | Taxonomy | OMIM | Books |

Nucleotide

Search [Nucleotide ▾] for [   ] [Go] [Clear]

Limits   Preview/Index   History   Clipboard   Details

[Display] [default ▾] [Save] [Text]   [Add to Clipboard]

☐ **1: U63139. Homo sapiens Rad5...**   ProbeSet, Related Sequences, OMIM, Protein, PubMed, Taxonomy,
[gi:1518805]   LinkOut

```
LOCUS       HSU63139               5892 bp    mRNA    linear   PRI 07-JUL-1999
DEFINITION  Homo sapiens Rad50 (Rad50) mRNA, complete cds.
ACCESSION   U63139
VERSION     U63139.1  GI:1518805
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 5892)
  AUTHORS   Dolganov,G.M., Maser,R.S., Novikov,A., Tosto,L., Chong,S.,
            Bressan,D.A. and Petrini,J.H.
  TITLE     Human Rad50 is physically associated with human Mre11:
            identification of a conserved multiprotein complex implicated in
            recombinational DNA repair
  JOURNAL   Mol. Cell. Biol. 16 (9), 4832-4841 (1996)
  MEDLINE   96347553
   PUBMED   8756642
REFERENCE   2  (bases 1 to 5892)
  AUTHORS   Dolganov,G.M., Maser,R.S., Novikov,A., Tosto,L., Chong,S.,
            Bressan,D.A. and Petrini,J.H.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-JUL-1996) Human Genetics, Genelabs Technologies,
            Inc., 505 Penobscot Drive, Redwood City, CA 94063, USA
FEATURES             Location/Qualifiers
     source          1..5892
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="5"
                     /map="5q31"
     gene            1..5892
                     /gene="Rad50"
     CDS             389..4327
                     /gene="Rad50"
                     /note="5'-end of mRNA is not verified by primer extension
                     or RNAse protection; the longest cDNA contains 388 bp of
                     5'UTR sequence"
                     /codon_start=1
                     /evidence=experimental
                     /product="Rad50"
                     /protein_id="AAB07119.1"
                     /db_xref="GI:1518806"
                     /translation="MSRIEKMSILGVRSFGIEDKDKQIITFFSPLTILVGPNGAGKTT
                     IIECLKYICTGDFPPGTKGNTFVHDPKVAQETDVRAQIRLQFRDVNGELIAVQRSMVC
                     TQKSKKTEFKTLEGVITRTKHGEKVSLSSKCAEIDREMISSLGVSKAVLNNVIFCHQE
                     DSNWPLSEGKALKQKFDEIFSATRYIKALETLRQVRQTQGQKVKEYQMELKYLKQYKE
                     KACEIRDQITSKEAQLTSSKEIVKSYENELDPLKNRLKEIEHNLSKIMKLDNEIKALD
                     SRKKQMEKDNSELEEKMEKVFQGTDEQLNDLYHNHQRTVREKERKLVDCHRELEKLNK
                     ESRLLNQEKSELLVEQCRLQLQADRHQEHIRARDSLIQSLATQLELDGFERGPFSERQ
```

```
IKNFHK    ERQEGEAKTANQLMNDFAEKETLKQKQIDE    KKTGLGRIIELKSEIL
SKKQNELKNVKYELQQLEGSSDRILELDQELIKAERELSKAEKNSNVETLKMEVISLQ
NEKADLDRTLRKLDQEMEQLNHHTTTRTQMEMLTKDKADKDEQIRKIKSRHSDELTSL
LGYFPNKKQLEDWLHSKSKEINQTRDRLAKLNKELASSEQNKNHINNELKRKEEQLSS
YEDKLFDVCGSQDFESDLDRLKEEIEKSSKQRAMLAGATAVYSQFTTQLTDENQSCCP
VCQRVFQTEAELQEVISDLQSKLRLAPDKLKSTESELKKKEKRRDEMLGLVPMRQSII
DLKEKEIPELRNKLQNVNRDIQRLKNDIEEQETLLGTIMPEEESAKVCLTDVTIMERF
QMELKDVERKIAQQAAKLQGIDLDRTVQQVNQEKQEKQHKLDTVSSKIELNRKLTQDQ
QEQIQHLKSTTNELKSEKLQISTNLQRRQQLEEQTVELSTEVQSLYREIKDAKEQVSP
LETTLEKFQQEKEELINKKNTSNKIAQDKLNDIKEKVKNIHGYMKDIENYIQDCKDDY
KKQKETELNKVIAQLSECEKHKEKINEDMRLMRQDIDTQKIQERWLQDNLTLRKRNEE
LKEVEEERKQHLKEMGQMQVLQMKSEHQKLEENIDNIKRNHNLALGRQKGYEEEIIHF
KKELREPQFRDAEEKYREMMIVMRTTELVNKDLDIYYKTLDQAIMKFHSMKMEEINKI
IRDLWRSTYRGQDIEYIEIRSDADENVSASDKRRNYNYRVVMLKGDTALDMRGRCSAG
QKVLASLIIRLALAETFCLNCGIIALDEPTTNLDRENIESLAHALVEIIKSRSQQRNF
QLLVITHDEDFVELLGRSEYVEKFYRIKKNIDQCSEIVKCSVSSLGFNVH"
```

BASE COUNT   2080 a   1097 c   1301 g   1414 t
ORIGIN
```
        1 ccaggagagc ggcgtggacg cgtgcgggcc tagaggccca cgtgatccgc agggcggccg
       61 aggcaggaag cttgtgagtg cgcggttgcg gggtcgcatt gtggctacgg ctttgcgtcc
      121 ccggcgggca gccccaggct ggtccccgcc tccgctctcc ccaccggcgg ggaaagcagc
      181 tggtgtggga ggaaaggctc catcccccgc cccctctctc ccgctgttgg ctggcaggat
      241 cttttggcag tcctgtggcc tcgctccccg cccggatcct cctgaccctg agattcgcgg
      301 gtctcacgtc ccgtgcacgc cttgcttcgg cctcagttaa gcctttgtgg actccaggtc
      361 cctggtgaga ttagaaacgt ttgcaaacat gtcccggatc gaaaagatga gcattctggg
      421 cgtgcggagt tttggaatag aggacaaaga taagcaaatt atcactttct tcagcccct
      481 tacaattttg gttggaccca atggggcggg aaagacgacc atcattgaat gtctaaaata
      541 tatttgtact ggagatttcc ctcctggaac caaaggaaat acatttgtac acgatcccaa
      601 ggttgctcaa gaaacagatg tgagagccca gattcgtctg caatttcgtg atgtcaatgg
      661 agaacttata gctgtgcaaa gatctatggt gtgtactcag aaaagcaaaa agacagaatt
      721 taaaactctg gaaggagtca ttactagaac aaagcatggt gaaaaggtca gtctgagctc
      781 taagtgtgca gaaattgacc gagaaatgat cagttctctt ggggtttcca aggctgtgct
      841 aaataatgtc attttctgtc atcaagaaga ttctaattgg ccttaagtg aaggaaaggc
      901 tttgaagcaa aagtttgatg agatttttc agcaacaaga tacattaaag ccttagaaac
      961 acttcggcag gtacgtcaga cacaaggtca gaaagtaaaa gaatatcaaa tggaactaaa
     1021 atatctgaag caatataagg aaaaagcttg tgagattcgt gatcagatta caagtaagga
     1081 agcccagtta acatcttcaa aggaaattgt caaatcctat gagaatgaac ttgatccatt
     1141 gaagaatcgt ctaaagaaa ttgaacataa tctctctaaa ataatgaaac ttgacaatga
     1201 aattaaagcc ttggatagcc gaaagaagca aatggagaaa gataatagtg aactggaaga
     1261 gaaaatggaa aaggttttc aagggactga tgagcaacta aatgacttat atcacaatca
     1321 ccagagaaca gtaaggagaa aagaaaggaa attggtagac tgtcatcgtg aactggaaaa
     1381 actaaataaa gaatctaggc ttctcaatca ggaaaaatca gaactgcttg ttgaacaggg
     1441 tcgtctacag ctgcaagcag atcgccatca agaacatatc cgagctagag attcattaat
     1501 tcagtctttg gcaacacagc tagaattgga tggctttgag cgtggaccat tcagtgaaag
     1561 acagattaaa aattttcaca aacttgtgag agagagacaa gaagggggaag caaaaactgc
     1621 caaccaactg atgaatgact ttgcagaaaa agagactctg aaacaaaaac agatagatga
     1681 gataagagat aagaaaactg gactgggaag aataattgag ttaaaatcag aaatcctaag
     1741 taagaagcag aatgagctga aaaatgtgaa gtatgaatta cagcagttgg aaggatcttc
     1801 agacaggatt cttgaactgg accaggagct cataaaagct gaacgtgagt taagcaaggc
     1861 tgagaaaaac agcaatgtag aaaccttaaa aatggaagta ataagtctcc aaaatgaaaa
     1921 agcagactta gacaggaccc tgcgtaaact tgaccaggag atggagcagt aaaccatca
     1981 tacaacaaca cgtacccaaa tggagatgct gaccaaagac aaagctgaca aagatgaaca
     2041 aatcagaaaa ataaaatcta ggcacagtga tgaattaacc tcactgttgg gatattttcc
     2101 caacaaaaaa cagcttgaag actggctaca tagtaaatca aaagaaatta atcagaccag
     2161 ggacagactt gccaaattga acaaggaact agcttcatct gagcagaata aaaatcatat
     2221 aaataatgaa ctaaaaagaa aggaagagca gttgtccagt tacgaagaca agctgtttga
     2281 tgtttgtggt agccaggatt ttgaaagtga tttagacagg cttaaagagg aaattgaaaa
     2341 atcatcaaaa cagcgagcca tgctggctgg agccacagca gtttactccc agttcattac
     2401 tcagctaaca gacgaaaacc agtcatgttg ccccgtttgt cagagagttt ttcagacaga
     2461 ggctgagtta caagaagtca tcagtgattt gcagtctaaa ctgcgacttg ctccagataa
     2521 actcaagtca acagaatcag agctaaaaaa aaaggaaaag cggcgtgatg aaatgctggg
     2581 acttgtgccc atgaggcaaa gcataattga tttgaaggag aaggaaatac cagaattaag
     2641 aaacaaactg cagaatgtca atagagacat acagcgccta aagaacgaca tagaagaaca
```

```
2701 agaaacactc ttgggt●    a taatgcctga agaagaaagt gccaaa     gcctgacaga
2761 tgttacaatt atggagaggt tccagatgga acttaaagat gttgaaagaa aaattgcaca
2821 acaagcagct aagctacaag gaatagactt agatcgaact gtccaacaag tcaaccagga
2881 gaaacaagag aaacagcaca agttagacac agtttctagt aagattgaat tgaatcgtaa
2941 gcttatacag gaccagcagg aacagattca acatctaaaa agtacaacaa atgagctaaa
3001 atctgagaaa cttcagatat ccactaattt gcaacgtcgt cagcaactgg aggagcagac
3061 tgtggaatta tccactgaag ttcagtcttt gtacagagag ataaaggatg ctaaagagca
3121 ggtaagccct ttggaaacaa cattggaaaa gttccagcaa gaaaaagaag aattaatcaa
3181 caaaaaaaat acaagcaaca aaatagcaca ggataaactg aatgatatta aagagaaggt
3241 taaaaatatt catggctata tgaaagacat tgagaattat attcaagatg ggaaagacga
3301 ctataagaag caaaaagaaa ctgaacttaa taaagtaata gctcaactaa gtgaatgcga
3361 gaaacacaaa gaaaagataa atgaagatat gagactcatg agacaagata ttgatacaca
3421 gaagatacaa gaaggtggc tacaagataa ccttacttta agaaaaagaa atgaggaact
3481 aaaagaagtt gaagaagaaa gaaaacaaca tttgaaggaa atgggtcaaa tgcaggtttt
3541 gcaaatgaaa agtgaacatc agaagttgga agagaacata gacaatataa aaagaaatca
3601 taatttggca ttagggcgac agaaaggtta tgaagaagaa attattcatt ttaagaaaga
3661 acttcgagaa ccacaatttc gggatgctga ggaaaagtat agagaaatga tgattgttat
3721 gaggacaaca gaacttgtga acaaggatct ggatatttat tataagactc ttgaccaagc
3781 aataatgaaa tttcacagta tgaaaatgga agaaatcaat aaaattatac gtgacctgtg
3841 gcgaagtacc tatcgtggac aagatattga atacatagaa atacggtctg atgccgatga
3901 aaatgtatca gcttctgata aaaggcggaa ttataactac cgagtggtga tgctgaaggg
3961 agacacagcc ttggatatgc gaggacgatg cagtgctgga caaaaggtat tagcctcact
4021 catcattcgc ctggccctgg ctgaaacgtt ctgcctcaac tgtggcatca ttgccttgga
4081 tgagccaaca acaaatcttg accgagaaa cattgaatct cttgcacatg ctctggttga
4141 gataataaaa agtcgctcac agcagcgtaa cttccagctt ctggtaatca ctcatgatga
4201 agatttgtg gagctttag gacgttctga atatgtggag aaattctaca ggattaaaaa
4261 gaacatcgat cagtgctcag agattgtgaa atgcagtgtt agctccctgg gattcaatgt
4321 tcattaaaaa tatccaagat ttaaatgcca tagaaatgta ggtcctcaga aagtgtataa
4381 taagaaactt attttctcata tcaacttagt caataagaaa atatattctt tcaaaggaac
4441 attgtgtcta ggattttgga tgttgagagg ttctaaaatc atgaaacttg tttcactgaa
4501 aattggacag attgcctgtt tctgatttgc tgctcttcat cccattccag gcagcctctg
4561 tcaggccttc agggttcagc agtacagccg agactcgact ctgtgcctcc ctcccagtg
4621 caaatgcatg cttcttctca aagcactgtt gagaaggaga taattactgc cttgaaaatt
4681 tatggttttg gtatttttt aaatcatagt taaatgttac ctctgaattt acttccttgc
4741 atgtggtttg aaaaactgag tattaatatc tgaggatgac cagaaatggt gagatgtatg
4801 tttggctctg cttttaactt tataaatcca gtgacctctc tctctgggac ttggtttccc
4861 caactaaaat ttgaagtagt tgaatggggt ctcaaagttt gacaggaacc ttaagtaatc
4921 atctaagtca gtacccacca ccttcttctc ctacatatcc cttccagatg gtcatccaga
4981 ctcagagctc tctctacaga gaggaaattc tccactgtgc acacccacct ttggaaagct
5041 ctgaccactt gaggcctgat ctgcccatcg tgaagaagcc tgtaacactc ctctgcgtct
5101 atcctgtgta gcatactggc ttcaccatca atcctgattc ctctctaagt gggcattgcc
5161 atgtggaagg caagccaggc tcactcacag agtcaaggcc tgctccctgt agggtccaac
5221 cagacctgga agaacaggcc tctccatttg ctcttcagat gccacttcta agaaaagcct
5281 aatcacagtt tttcctggaa ttgccagctg acatcttgaa tccttccatt ccacacagaa
5341 tgcaaccaag tcacacgctt ttgaattatg ctttgtagag ttttgtcatt cagagtcagc
5401 caggaccata ccgggtcttg attcagtcac atggcatggt tttgtgccat ctgtagctat
5461 aatgagcatg tttgcctaga cagctttct caactgggtc cagaagagaa ttaagcccta
5521 aggtcctaag gcatctatct gtgctaggtt aaatggttgg cccccaaaga tagacaggtc
5581 ctgatttcta gaacccgtga ctgttacttt atacagcaaa ggaaactttg cagatgtgat
5641 taaagctaag gaccttaaga cagagtatcc tggggtggt ggtgggtgg gggggggtcc
5701 taaatgtaat cacgagtaag attaagagca aatcaattct agtcatatat taaacatcca
5761 caataaccaa gatattttta tcccaagaat gcaagatttc agaaaatgaa aaatctgttg
5821 ataaatccat cactataata aaaccgaagg tgaaaaaat tctgaaaaa aaaaaaaaaa
5881 aaaaaaaaaa aa
//
```

Revised: October 24, 2001.

NCBI

Nucleotide
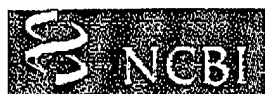
PubMed   Nucleotide   Protein   Genome   Structure   PopSet   Taxonomy   OMIM   Books

Search | Nucleotide ▾ | for | | Go | Clear |

Limits   Preview/Index   History   Clipboard   Details

| Display | default ▾ | Save | Text | Add to Clipboard |

☐ 1: Z75311. H.sapiens mRNA fo...[gi:2687852]   Related Sequences, OMIM, Protein, Taxonomy, LinkOut

```
LOCUS       HSRAD50                 4123 bp    mRNA    linear   PRI 11-DEC-1997
DEFINITION  H.sapiens mRNA for RAD50.
ACCESSION   Z75311
VERSION     Z75311.1  GI:2687852
KEYWORDS    RAD50.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 4123)
  AUTHORS   Offenberg,H.H.
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 4123)
  AUTHORS   Offenberg,H.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-JUL-1996) Offenberg H.H., Agricultural University,
            Genetics, Dreyenlaan 2, Wageningen, 6703 HA The Netherlands
FEATURES             Location/Qualifiers
     source          1..4123
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /tissue_type="testis"
                     /clone_lib="library HL1161a in lambda gt10 (Clontech)"
     gene            43..3999
                     /gene="RAD50"
     CDS             43..3999
                     /gene="RAD50"
                     /function="DNA repair and recombindation protein"
                     /codon_start=1
                     /product="RAD50 homologue hsRAD50"
                     /protein_id="CAA99729.1"
                     /db_xref="GI:2687853"
                     /db_xref="SPTREMBL:O43254"
                     /translation="MLIFSVRDMFAKMSILGVRSFGIEDKDKQIITFFSPLTILVGPN
                     GAGKTTIIECLKYICTGDFPPGTKGNTFVHDPKVAQETDVRAQIRLQFRDVNGELIAV
                     QRSMVCTQKSKKTEFKTLEGVITRTKHGEKVSLSSKCAEIDREMISSLGVSKAVLNNV
                     IFCHQEDSNWPLSEGKALKQKFDEIFSATRYIKALETLRQVRQTQGQKVEEYQMELKY
                     LKQYKEKACEIRDQITSKEAQLTSSKEIVKSYENELDPLKNRLKEIEHNLSKIMKLDN
                     EIKALDSRKKQMEKDNSELEEKMEKVFQGTDEQLNDLYHNHQRTVREKERKLVDCHRE
                     LEKLNKESRLLNQEKSELLVEQGRLQLQADRHQEHIRARDSLIQSLATQLELDGFERG
                     PFSERQIKNFHKLVRERQEGEAKTANQLMNDFAEKETLKQKQIDEIRDKKTGLGRIIE
                     LKSEILSKKQNELKNVKYELQQLEGSSDRILELDQELIKAERELSKAEKNSNVETLKM
                     EVISLQNEKADLDRTLRKLDQEMEQLNHHTTTRTQMEMLTKDKADKDEQIRKIKSRHS
                     DELTSLLGYFPNKKQLEDWLHSKSKEINQTRDRLAKLNKELASSEQNKNIINNELERK
                     EEQLSSYEDKLFDVCGSQDFESDLDRLKEEIEKSSKQRAMLAGATAVYSQFITQLTDE
                     NQSCCPVCQRVFQTEAELQEAISDLQSKLRLAPDKLKSTESELKKKEKRRDEMLGLAP
                     MRQSIIDLKEKEIPELRNKLQNVNRDIQRLKNDIEEQETLLGTIMPEEESAKVCLTDV
                     TIMERFQMELKDVERKIAQQAAKLQGIDLDRTVQQVNQEKQEKQHKLDTVSSKIELNR
                     KLIQDQQEQIQHLKSTTNELKSEKLQISTNLQRRQQLEEQTVELSTEVQSLYREIKDA
                     KEQVSPLETTLEKFQQEKEELINKKNTSNKIAQDKLNDIKEKVKNIHGYMKDIENHIQ
```

```
                 DGKDDY●KETELNKVIAQLSECEKHKEKINEDMRLMR●DTQKIQERWLQDNLTL
                 RKRNEELKEVEEEGKQHLKEMGQMQVLQMKSEHQKLEENIDNIKRNHNIAIGRQKGYE
                 EEIIHFKKELREPQFRDAEEKYREMMIVMRTTELVNKDLDIYYKTLDQAIMKFHSMKM
                 EEINKIIRDLWRSTYRGQDIEYIEIRSDADENVSASDKRRNYNYRVVMLKGDTALDMR
                 GRCSAGQKVLASLIIRLALAETFCLNCGIIALDEPTTNLDRENIESLAHALVEIIKSR
                 SQQRNFQLLVITHDEDFVELLGRSEYVEKFYRIKKNIDQCSEIVKCSVSSLGFNVH"
```

BASE COUNT   1630 a   661 c   901 g   931 t
ORIGIN

```
   1 tgcggagttt tggaatagag gacaaagata agcacgccca gaatgctcat cttttcggtc
  61 cgggacatgt ttgcaaagat gagcattctg ggcgtgcgga gttttggaat agaggacaaa
 121 gataagcaaa ttatcacttt cttcagcccc cttacaattt tggttggacc caatggggcg
 181 ggaaagacga ccatcattga atgtctaaaa tatatttgta ctggagattt ccctcctgga
 241 accaaaggaa atacatttgt acacgatccc aaggttgctc aagaaacaga tgtgagagcc
 301 cagattcgtc tgcaatttcg tgatgtcaat ggagaactta tagctgtgca aagatctatg
 361 gtgtgtactc agaaaagcaa aaagacagaa tttaaaaccc tggaaggagt cattactaga
 421 acaaagcatg gtgaaaaggt cagtctgagc tctaagtgtg cagaaattga ccgagaaatg
 481 atcagttctc ttggggtttc caaggctgtg ctaaataatg tcattttctg tcatcaagaa
 541 gattctaatt ggcctttaag tgagggaaag gctttgaagc aaaagtttga tgagattttt
 601 tcagcaacaa gatacattaa agccctagaa acacttcggc aggtacgtca gacacaaggt
 661 cagaaagtag aagaatatca aatggaacta aaatatctga agcaatataa ggaaaaagct
 721 tgtgagattc gtgatcagat tacaagtaag gaagcccagt taacatcttc aaaggaaatt
 781 gtcaaatcct atgagaatga acttgatcca ttgaagaatc gtctaaaaga aattgaacat
 841 aatctctcta aaataatgaa acttgacaat gaaattaaag ccttggatag ccgaaagaag
 901 caaatggaga aagataatag tgaactggaa gagaaaatgg aaaaggtttt tcaagggact
 961 gatgagcaac taaatgactt atatcacaat caccagagaa cagtaaggga gaaagaaagg
1021 aaattggtag actgtcatcg tgaactggaa aaactaaata aagaatctag gcttctcaat
1081 caggaaaaat cagaactgct tgttgaacag ggtcgtctac agctgcaagc agatcgccat
1141 caagaacata tccgagctag agattcatta attcagtctt tggcaacaca gctagaattg
1201 gatggctttg agcgtggacc attcagtgaa agacagatta aaaattttca caaacttgtg
1261 agagagagac aagaagggga agcaaaaact gccaaccaac tgatgaatga ctttgcagaa
1321 aaagagactc tgaaacaaaa acagatagat gagataagag ataagaaaac tggactggga
1381 agaataattg agttaaaatc agaaatccta agtaagaagc agaatgagct gaaaaatgtg
1441 aagtatgaat tacagcagtt ggaaggatct tcagacagga ttcttgaact ggaccaggag
1501 ctcataaaag ctgaacgtga gttaagcaag gctgagaaaa acagcaatgt agaaacctta
1561 aaaatggaag taataagtct ccaaaatgaa aaagcagact tagacaggac cctgcgtaaa
1621 cttgaccagg agatggagca gttaaaccat catacaacaa cacgtacccca aatggagatg
1681 ctgaccaaag acaaagctga caaagatgaa caaatcagaa aaataaaatc taggcacagt
1741 gatgaattaa cctcactgtt gggatatttt cccaacaaaa aacagcttga agactggcta
1801 catagtaaat caaaagaaat taatcagacc agggacagac ttgccaaatt gaacaaggaa
1861 ctagcttcat ctgagcagaa taaaaatcat ataaataatg aactagaaag aaaggaagag
1921 cagttgtcca gttacgaaga caagctgttt gatgtttgtg gtagccagga ttttgaaagt
1981 gatttagaca ggcttaaaga ggaaattgaa aaatcatcaa aacagcgagc catgctggct
2041 ggagccacag cagtttactc ccagttcatt actcagctaa cagacgaaaa ccagtcatgt
2101 tgccccgttt gtcagagagt tttttcagaca gaggctgagt tacaagaagc catcagtgat
2161 ttgcagtcta aactgcgact tgctccagat aaactcaagt caacagaatc agagctaaaa
2221 aaaaaggaaa agcggcgtga tgaaatgctg ggacttgcgc ccatgaggca aagcataatt
2281 gattt gaagg agaaggaaat accagaatta agaaacaaac tgcagaatgt caatagagac
2341 atacagcgcc taaagaacga catagaagaa caagaaacac tcttgggtac aataatgcct
2401 gaagaagaaa gtgccaaagt atgcctgaca gatgttacaa ttatggagag gttccagatg
2461 gaacttaaag atgttgaaag aaaaattgca caacaagcag ctaagctaca aggaatagac
2521 ttagatcgaa ctgtccaaca agtcaaccag gagaaacaag agaaacagca caagttagac
2581 acagtttcta gtaagattga attgaatcgt aagcttatac aggaccagca ggaacagatt
2641 caacatctaa aaagtacaac aaatgagcta aaatctgaga aacttcagat atccactaat
2701 ttgcaacgtc gtcagcaact ggaggagcag actgtggaat tatccactga agttcagtct
2761 ttgtacagag agataaagga tgctaaagag caggtaagcc ctttggaaac aacattggaa
2821 aagttccagc aagaaaaaga agaattaatc aacaaaaaa atacaagcaa caaaatagca
2881 caggataaac tgaatgatat taaagagaag gttaaaaata ttcatggcta tatgaaagac
2941 attgagaatc atattcaaga tgggaaagac gactatatga agcaaaaaga aactgaactt
3001 aataaagtaa tagctcaact aagtgaatgc gagaaacaca agaaaagat aaatgaagat
3061 atgagactca tgagacaaga tattgataca cagaagatac aagaaggtg gctacaagat
3121 aaccttactt taagaaaaag aaatgaggaa ctaaaagaag ttgaagaaga aggaaaacaa
3181 catttgaagg aaatgggtca aatgcaggtt ttgcaaatga aaagtgaaca tcagaagttg
3241 gaagagaaca tagacaatat aaaaagaaat cataatttgg cattagggcg acagaaaggt
```

```
3301  tatgaagaag aaatta███a ttttaagaaa gaacttcgag aaccac███t tcgggatgct
3361  gaggaaaagt atagagaaat gatgattgtt atgaggacaa cagaacttgt gaacaaggat
3421  ctggatattt attataagac tcttgaccaa gcaataatga aatttcacag tatgaagatg
3481  gaagaaatca ataaaattat acgtgacctg tggcgaagta cctatcgtgg acaagatatt
3541  gaatacatag aaatacggtc tgatgccgat gaaaatgtat cagcttctga taaaaggcgg
3601  aattataact accgagtggt gatgctgaag ggagacacag ccttggatat gcgaggacga
3661  tgcagtgctg gacaaaaggt attagcctca ctcatcattc gcctggccct ggctgaaacg
3721  ttctgcctca actgtggcat cattgccttg gatgagccaa caacaaatct tgaccgagaa
3781  aacattgaat ctcttgcaca tgctctggtt gagataataa aaagtcgctc acagcagcgt
3841  aacttccagc ttctggtaat cactcatgat gaagattttg tggagctttt aggacgttct
3901  gaatatgtgg agaaattcta caggattaaa aagaacatcg atcagtgctc agagattgtg
3961  aaatgcagtg ttagctccct gggattcaat gttcattaaa aatatccaag atttaaatgc
4021  catagaaatg taggtcctca gaaagtgtat aataagaaac ttatttctca tatcaactta
4081  gtcaataaga aaatatattc tttcaaagga aaaaaaaaaa aaa
```

//

Revised: October 24, 2001.

NCBI   **Nucleotide**

PubMed   Nucleotide   Protein   Genome   Structure   PopSet   Taxonomy   OMIM   Books

Search [Nucleotide ▼] for [ ]   [Go] [Clear]

Limits   Preview/Index   History   Clipboard   Details

[Display] [default ▼] [Save] [Text]   [Add to Clipboard]

☐ **1: NM_022246. Rattus norvegicus...[gi:11560047]**   Related Sequences, Protein, PubMed, Taxonomy, LinkOut

```
LOCUS       NM_022246               4444 bp    mRNA    linear   ROD 06-DEC-2000
DEFINITION  Rattus norvegicus RAD50 homolog (S. cerevisiae) (Rad50), mRNA.
ACCESSION   NM_022246
VERSION     NM_022246.1  GI:11560047
KEYWORDS    .
SOURCE      Norway rat.
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 4444)
  AUTHORS   Lanson,N.A. Jr., Egeland,D.B., Royals,B.A. and Claycomb,W.C.
  TITLE     The MRE11-NBS1-RAD50 pathway is perturbed in SV40 large T
            antigen-immortalized AT-1, AT-2 and HL-1 cardiomyocytes
  JOURNAL   Nucleic Acids Res. 28 (15), 2882-2892 (2000)
  MEDLINE   20368653
COMMENT     PROVISIONAL REFSEQ: This record has not yet been subject to final
            NCBI review. The reference sequence was derived from AF218576.1.
FEATURES             Location/Qualifiers
     source          1..4444
                     /organism="Rattus norvegicus"
                     /db_xref="taxon:10116"
     gene            1..4444
                     /gene="Rad50"
                     /db_xref="LocusID:64012"
     CDS             156..4094
                     /gene="Rad50"
                     /function="DNA double-strand break repair and
                     recombination"
                     /function="telomere length maintenance"
                     /note="similar to Saccharomyces cerevisiae Rad50; null
                     mutation is lethal in murine embryonic stem cells"
                     /codon_start=1
                     /db_xref="LocusID:64012"
                     /product="RAD50 homolog (S. cerevisiae)"
                     /protein_id="NP_071582.1"
                     /db_xref="GI:11560048"
                     /translation="MSRIEKMSTLGVRSFGIEDKDKQIISFFSPLTILVGPNGAGKTT
                     IIECLKYICTGDFPPGTKGNTFVHDPKVAQETDVRAQIRLQFRDVNGEMVLVQRSMLC
                     SQKSKKTEFKTLEGVITRIKHGEKVSLSSKCAEIDREMISCLGVSKSVLNNVIFCHQE
                     DSNWPLSEGKALKQKFDEIFSATRYIKALDTLRQVRQTQGQKVKECQTELKYLRQNKE
                     KACEIRDQITSKEAQLASSRETVKAYENELEPLKNRLKEIEHNLSKIMRLDNEIKALD
                     SRKKQMEKDNSELEQKMEKVFQGTDEQLNDLYHNHQRTVREKERRLVDCQRELEKLSK
                     EARLLNQERAELLVEQGRLQLQADRHQEHIRARDSLIQSLAAHLELDGFERGPFSERQ
                     IKNFHELVRERQEREAKTASQLLSDLTDKEALKQRQMDEMRDKKSGLGRMIELKTEIL
                     TKKQTELRNVRNELQQLEGSSDRILELDQELTKAERELSKAEKNSSIETLKAEILNLQ
                     SEKADLDRNLRKLDQEMEQLNHHTTTRTQMEMLTKDKTDKDEQIRKIKSRHSDELTSL
                     LGYFPNKKQLEDWLHSKSKEINQTRDRLAKLNKELASAEQNKNHINNELKKKEEQLSS
                     YEDKLFDVCGSQDFESDLDRLKEDIEKSSKQRAMLAGATAVYSQFITQLTDENQSCCP
                     GCQRVFQTEAELQEVISDLQSKLRLAPDKLKSTESELKKKERRRDEMLGLVPMRQSII
```

DLKEKE RNRLQSVNRDIQRLKNDIEEQETLLGTVM SAKVCLTDVTIMERF
QMELKDVERKIAQQAAKLQGVDLDRTVQQVNQEKQEKQHKLDTVSSKIELNRKLIQDQ
QEQIQHLKSKTNELKSEKLQIATNLQRRQQMEEQTVELSTEVQSLNREIKDAKEQINP
LEIALEKLQQEKEELIHRKNTSNKMAQDKINDIKEKVKNIHGYMKDIENYIQDGKDDY
KKQKETELNEVVIQLNECDKHKEKINKEMGTMRODIDTKKIQERWLQDNLTLRKRREE
LKEVEEERKQHLKEMGQMQVLQMKNEHQKLEENIDTIKRNHSLALCRQKGYEEEILHF
KKELREPQFRDAEEKYREMMIVMRTTELVNKDLDIYYKTLDHAIMKFHSMKMEEINKI
TRDLWRSTYRGQDIEYIEIRSDADENVSASDKRRNYNYRVVMLKGDTALDMRGRCSAG
QKVLASLIIRLALAETFCLNCGILALDEPTTNLDRENIESLAHALVEIIKSRSQQRNF
QLLVITHDEDFVELLGRSEYVEKFYRVKKNIDQCSEIVKSSTNSLGSYVH"

misc feature	846..2039
	/note="Myosin_tail; Region: Myosin tail"
misc feature	990..1958
	/note="filament; Region: Intermediate filament proteins"
misc feature	1635..2144
	/note="V_ATPase_sub_a; Region: V-type ATPase 116kDa
	subunit family"
misc feature	3705..3986
	/note="ABC_tran; Region: ABC transporter"

BASE COUNT   1549 a   840 c   1116 g   939 t
ORIGIN
```
   1 tgcgggccg gaagtgctct cgccgccgga gctcccactc gcagctcggg ttcggaactt
  61 tcgggtcctc tcgtggcgtc ccttcccggc ttgcccagct tccgcttcga ggttaaactg
 121 gcccctggca gtgagataag aagcgtggtg caaacatgtc ccggatcgaa aagatgagca
 181 ctctgggcgt gcgaagtttt gggatagagg ataaagataa gcaaattatc tctttcttca
 241 gcccccncac aattttggtt ggacccaatg gggcggggaa gacgaccatc attgaatgtc
 301 taaagtatat ttgtactgga gatttccctc ctggaaccaa aggaaataca tttgttcatg
 361 atcccaaggt tgctcaagaa acagatgtgc gtgcccaaat tcgcctgcag tttcgtgatg
 421 tcaatggaga gatggtactt gtgcagaggt ccatgctttg cagtcagaaa agtaaaaaaa
 481 ccgaatttaa aaccctggaa ggagtcatta ctagaataaa gcacggtgaa aaagtcagtc
 541 tcagctccaa atgtgcagaa atcgaccgag aaatgataag ttgtcttggg gtttccaagt
 601 ctgtgctaaa caatgttatt ttctgccacc aagaagactc aaattggcct ctaagtgaag
 661 gaaaggctct gaagcagaaa tttgatgaga tttttttcagc aacaaggtac attaaagccc
 721 tagatacgct tcgacaggta cgacagacac aaggtcagaa agtaaaagaa tgtcaaacag
 781 aattgaaata tctgaggcaa aataaggaga aagcttgtga gatccgagat cagatcacta
 841 gtaaggaagc ccagttagca tcttcacggg aaattgtcaa agcctatgag aatgagcttg
 901 agccattgaa gaatcgcctg aaagagattg aacataacct ctctaaaata atgagacttg
 961 acaatgaaat taaagccttg gatagcagaa agaagcaaat ggaaaaagat aacagtgaat
1021 tagaacagaa gatggaaaag gtttttcaag ggactgatga gcagctaaat gacttgtatc
1081 acaatcacca gagaactgta agggagaaag aaaggcgctt ggtagactgt cagcgtgaac
1141 tggagaagct gagtaaagaa gctcggctcc tcaaccagga aagagcagag ctgcttgtgg
1201 agcagggtcg tctacaacta caggcagatc gacatcaaga gcatatccga gccagagact
1261 cactgattca gtctttggca gcacatcttg aattggatgg ttttgagcgt ggaccattca
1321 gtgaaagaca gattaaaaac tttcatgaac ttgtgagaga gagacaggag agagaagcta
1381 aaactgccag ccagctcttg agcgacctta cagacaaaga agcggctgaag cagagacaga
1441 tggatgagat gagggacaag aagagcgggc tggggaggat gattgagctg aagaccgaga
1501 tcctgacgaa gaagcagact gagctgagga acgtgaggaa tgagctgcag cagctggagg
1561 gctcctcgga caggattctg gagctggacc aggagctcac aaaagcggaa cgtgaactaa
1621 gcaaggctga gaaaaatagc agcatagaaa ccctaaaagc agaaatacta aacctccaaa
1681 gtgagaaagc ggacctggac aggaacctgc ggaaactgga tcaggagatg gagcagttaa
1741 accatcatac aacaacccgc acacagatgg agatgcttac caaagacaaa actgacaaag
1801 atgaacagat cagaaaaata aagtccaggc acagtgatga actaacttca ctgttgggat
1861 attttcctaa caaaaaacag cttgaagact ggcttcattc taaatccaaa gagattaatc
1921 agaccaggga cagacttgcc aaactgaaca agaactagc ttcagccgaa caaaataaaa
1981 atcatataaa taatgagcta aagaaaaagg aagagcagct gtctagttat gaagataaac
2041 tgtttgatgt ttgtggtagc caagattttg aaagtgactt agacagactt aaagaagata
2101 ttgaaaaatc ctcaaagcag cgagccatgc tggctggagc cacagcagtt tactcccagt
2161 tcatcactca gctgacagat gaaaaccagt cctgttgccc aggctgtcag agagtatttc
2221 agacggaagc tgaattacag gaagtcatca gtgacttgca gtccaagctg aggcttgctc
2281 cagataaact caagtcaaca gaatcagaac taaaaaaaaa agagcggcgc cgtgatgaaa
2341 tgctgggggct tgtgcccatg aggcaaagca taattgattt gaaggaaaag gaaataccag
2401 aattaagaaa cagactgcag agtgtcaata gagacataca gcgcctaaag aatgacatag
2461 aggagcagga gacactcttg ggtacagtga tgcccgaaga ggaaagtgct aaagtgtgcc
```

```
2521 tgacagacgt cacgat●●g gagaggttcc agatggagct aaaaga●●c gaaaggaaaa
2581 ttgcacagca ggcagctaag ctgcaggggg tagacttgga tcggactgtc cagcaggtta
2641 accaggaaaa acaagaaaaa caacacaaac tggatacagt ttccagtaag attgaattga
2701 accgtaagct tatacaggac cagcaggaac aaattcagca cctgaaaagt aaaacaaatg
2761 agctgaaatc agagaaactg cagatagcca ccaatttgca acggcgtcag caaatggagg
2821 agcagactgt ggaattatcc actgaagttc agtctttgaa cagagaaata aaggatgcta
2881 aagagcaaat aaaccctttg gagatagcac tggaaaagtt gcagcaggaa aaagaagaat
2941 tgatccacag aaaaaataca agtaacaaaa tggctcagga taagatcaat gatatcaaag
3001 agaaagttaa aaatattcat ggttacatga aagacataga gaattatatt caagatggaa
3061 aagatgacta taagaagcaa aaagaaactg aacttaatga agttgtcatt caactaaatg
3121 aatgtgacaa acacaaagaa aagataaata aagaaatggg aaccatgagg caagatattg
3181 acacgaagaa gatacaggaa aggtggttac aggataacct tacactgaga aaaagaagag
3241 aggaactaaa agaagttgaa gaggaacgaa agcaacattt gaaggagatg ggacagatgc
3301 aagtttttaca gatgaaaaat gagcaccaga agttggaaga gaacatagac acgataaaga
3361 gaaatcacag tttggcattg gggcgacaga aaggctatga agaagaaatt cttcacttta
3421 agaaggagct ccgagaacct cagtttcggg atgcagagga gaagtacaga gaaatgatga
3481 ttgtcatgag aaccacagag ctggtgaata aggacctgga catctactac aaaactcttg
3541 atcacgcaat aatgaaattt cacagtatga aaatggaaga aatcaacaaa attattcgtg
3601 atctttggcg gagtacctat cgtgggcaag atattgaata catagaaatt cgatccgatg
3661 ctgatgaaaa tgtatcagct tctgataaaa ggcgaaatta caactaccga gtggtgatgc
3721 tgaaggggga tacagccttg gacatgcgcg gacgatgcag tgctggacag aaggtgctgg
3781 cctctctcat catccgactg gccctggccg aaaccttctg tctgaactgc ggcatccttg
3841 ccttggatga gcctacaaca aacctggacc gagaaaacat cgagtctctg gcacatgctt
3901 tggttgagat tataaaaagt cgctcgcagc agcgcaactt ccagcttctg gtcatcactc
3961 acgatgaaga ttttgtggag ctcctaggac gatctgagta tgtggagaaa ttctacagag
4021 tgaagaagaa catcgaccag tgctcagaga ttgtcaagag cagcatcaac tctctgggct
4081 cttatgttca ctgacttgtt ttgggaggga tgtggattgg gtgaagggaa ctgtgtatgg
4141 caggtttgag agttcctgag atcgtgaggt tgccttgaat gaagactgaa cacagttccc
4201 acaggcctag caacctccca gttctgctca gctgtttatt tcctgcctac tgaatgtcac
4261 ccagataaac agtgctcttg agagcagtgc tgacaggagt ggtgagcatt tcaaagtgtt
4321 gctgttttct ctgaagctta gagtgatatt tagtaagtgg ccctgcttta acatgaatcc
4381 atgcacacaa cctgtctgtc cgagcctagg tcttccctga aaccacagtc attgaacggg
4441 gcta
//
```

┌─────────┐
│ ⬡ NCBI │   [navigation icons]   Nucleotide
└─────────┘

PubMed   Nucleotide   Protein   Genome   Structure   PopSet   Taxonomy   OMIM   Books

Search [Nucleotide ▼] for [_____]   [Go] [Clear]

            Limits   Preview/Index   History   Clipboard   Details

[Display] [default ▼] [Save] [Text]   [Add to Clipboard]

☐ 1: U66887. Mus musculus DNA ...[gi:1575574]   ProbeSet, Related Sequences, Protein, PubMed, Taxonomy, LinkOut

```
LOCUS       MMU66887                5088 bp    mRNA    linear   ROD 15-NOV-1996
DEFINITION  Mus musculus DNA repair protein RAD50 (RAD50) mRNA, complete cds.
ACCESSION   U66887
VERSION     U66887.1  GI:1575574
KEYWORDS    .
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 5088)
  AUTHORS   Kim,K.K., Daud,A.I., Wong,S.C., Pajak,L., Tsai,S.C., Wang,H.,
            Henzel,W.J. and Field,L.J.
  TITLE     Mouse RAD50 has limited epitopic homology to p53 and is expressed
            in the adult myocardium
  JOURNAL   J. Biol. Chem. 271 (46), 29255-29264 (1996)
  MEDLINE   97067183
REFERENCE   2  (bases 1 to 5088)
  AUTHORS   Kim,K.K., Daud,A.I., Wong,S.C., Pajak,L., Tsai,S.C., Wang,H.,
            Henzel,W.J. and Field,L.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-AUG-1996) Medicine, Indiana University, Krannert
            Institute of Cardiology, 1111 West 10th Street, Indianapolis, IN
            46202-4800, USA
FEATURES             Location/Qualifiers
     source          1..5088
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /chromosome="11"
                     /map="A5-B1Mou"
                     /cell_type="AT-2 cardiomyocytes"
     gene            1..5088
                     /gene="RAD50"
     CDS             243..4181
                     /gene="RAD50"
                     /function="DNA repair protein"
                     /codon_start=1
                     /product="RAD50"
                     /protein_id="AAC52894.1"
                     /db_xref="GI:1575575"
                     /translation="MSRIEKMSILGVRSFGIEDKDKQIISFFSPLTILVGPNGAGKTT
                     IIECLKYICTCDFPPGTKGNTFVHDPKVAQETDVRAQIRLQFRDVNGEMVAVHRSMLC
                     SQKNKKTEFKTLEGVITRMKHGEKVSLSSKCAEIDREMISCLGVSKSVLNNVIFCHQE
                     DSNWPLSEGKALKQKFDEIFSATRYIKALDTLRQVRQTQGQKVKECQTELKYLKQNKE
                     KACETRDQITSKEAQLASSQEIVRSYEDELEPLKNRLKETEHNLSKTMKLDNEIKALE
                     SRKKQMEKDNSELEQKMEKVFQGTDEQLNDLYIINIIQRTVREKERRLVDCQRELEKLNK
                     EARLNQEKAELLVEQGRLQLQADRHQEHTRARDSLIQSLATHLELDGFERGPFSERQ
                     IKNFHELVKERQEREAKTASQLLSDLTDKEALKQRQLDELRDRKSGLGRTIELKTEIL
                     TKKQSELRHVRSELQQLEGSSDRILELDQELTKAERELSKAEKNSSIETLKAEVMSLQ
                     NEKADLDRSLRKLDQEMEQLNHHTTTRTQMEMLTKDKTDKDEQIRKIKSRHSDELTSL
                     LCYFPNKKQLEDWLHSKSKEINQTRDRLAKLNKELASAEQNKNHINNELKKKEEQLSS
```

```
        YEDKLF●●SQDLESDLGRLKEEIEKSSKQRAMLAGAT●●QFITQLTDENQSCCP
        VCQRVFQTEAELQEVISDLQSKLRLAPDKLKSTESELKKKERRRDEMLGLVPVRQSII
        DLKEKEIPELRNRLQSVNRDIQRLKNDIEEQETLLGTIMPEEESAKVCLTDVTIMERF
        QMELKDVERKIAQQAAKLQGVDLDRTVQQVNQEKQEKQHRLDTVTSKIELNRKLIQDQ
        QEQIQHLKSKTNELKSEKLQIATNLQRRQQMEEQSVELSTEVQSLNREIKDAKEQISP
        LETALEKLQQEKEELIHRKHTSNKMAQDKINDIKEKVKNIHGYMKDIENYIQDGKDDY
        KKQKETELNGVAVQLNECEKHREKINKDMGTMRQDIDTQKIQERWLQDNLTLRKRRDE
        LKEVEEEPKQHLKEMGQMQVLQMKNEHQKLEENIDTIKRNHSLALGRQKGYEDEILHF
        KKELREPQFRDAEEKYREMMIVMRTTELVNKDLDIYYKTLDQAIMKFHSMKMEEINKI
        IRDLWRSTYRGQDIEYIEIRSDADENVSASDKRRNYNYRVVMLKGDTALDMRGRCSAG
        QKVLASLITRLALAETFCLNCGILALDEPTTNLDRENIESLAHALVETIKSRSQQRNF
        QLLVITHDEDFVELLGRSEYVEKFYRVKKNMDQCSEIVKCSISSLGSYVH"
BASE COUNT     1705 a    1036 c    1264 g    1083 t
ORIGIN
        1 gtgcgagctt acggttgctg gctcgcttgg cttttctggg ctgtgcagac cctaaccggc
       61 ccgcgccctc ccgggaaacc gaagcccgcg gggccggaag tgctctcgcc accggagctc
      121 ctcgtgccag cggctgttcg gaacttttgg gtcctctcgc ggcgttcctt cccggcttgc
      181 ccagcttccg ccttgaggtt aaaactcgcc cctggcactg agtaaaggaa cgtggtgcaa
      241 acatgtcccg gatcgaaaag atgagcattc tgggcgtgcg aagttttggg atagaggata
      301 aagataagca gattatctct ttcttcagcc ccctcacaat tttggttgga cccaatgggg
      361 cggggaagac gaccatcatt gaatgtctaa agtatatttg tactggagat ttccctcctg
      421 gaaccaaagg aaatacattt gttcatgatc ccaaggttgc tcaagaaaca gacgtgcggg
      481 cccagattcg cctgcagttt cgagatgtga atggagagat ggtagctgtg cacaggtcta
      541 tgctttgcag tcagaaaaat aaaaaaacag aatttaaaac cttggaagga gtcattacta
      601 gaatgaagca tggtgaaaaa gtcagtctca gctccaaatg tgcagaaatc gacagagaaa
      661 tgataagttg tcttgggtt tccaagtctg tgctaaacaa tgttatttc tgtcaccaag
      721 aagactccaa ctggccttta agtgaaggaa aggctttgaa gcaaaaattt gatgagattt
      781 tttcggcaac aaggtacatt aaagccttag atacacttcg acaggtgcgt cagacacagg
      841 gtcagaaagt aaaagagtgt caaacggaat tgaaatatct gaagcaaaat aaggagaaag
      901 cgtgtgagat ccgagatcag attacgagta aggaagccca gttagcgtct tcgcaggaaa
      961 ttgtcaggtc ctacgaggat gagcttgagc cattgaagaa tcgcctgaaa gagattgaac
     1021 acaacctgtc taaaataatg aaacttgata atgaaattaa agccttggag agccgaaaga
     1081 aacaaatgga gaaagataac agtgaattag aacagaaaat ggaaaaggtt tttcaaggga
     1141 ctgatgagca gctaaatgac ttgtatcaca atcaccagag aacagtaagg gagaaagaaa
     1201 ggcgcctggt agactgccag cgagaactgg agaagctgaa taaagaagca cggctcctca
     1261 accaggaaaa agcagagctg cttgtggagc agggtcgtct acagctacag gcagatcgcc
     1321 atcaagagca catccgagcc agagactcat taattcagtc tttggcaaca catcttgaat
     1381 tggatggttt tgagcgtgga ccattcagtg aaagacagat taaaaatttt catgaacttg
     1441 tgaaagagag acaggaaaga gaagctaaaa ctgccagcca gctcttgagc gacctgacag
     1501 acaaagaagc tctgaagcag aggcagcttg atgagctgag ggacaggaag agcggactgg
     1561 ggaggacgat tgagctgaag acagagatcc tcacgaagaa gcagagcgag ctgcggcacg
     1621 tgaggagtga gctgcagcag ctggagggct cctccgacag gattctggaa ctggaccagg
     1681 agctcacgaa agctgaacgt gaattaagca aggccgagaa aaatagcagc atagaaaccc
     1741 taaaagcaga ggtaatgagc ctccaaaatg agaaagctga cctggacagg agcttgcgaa
     1801 agctggacca ggagatggag cagttaaacc atcatacaac aacccgcaca cagatggaga
     1861 tgctcaccaa agacaaaact gacaaagatg aacagatcag aaaaataaag tctaggcaca
     1921 gtgatgaact aacttcactg ttggggtact ttcctaacaa aaaacagctt gaagactggc
     1981 ttcattctaa atccaaagaa attaatcaga ccagggacag acttgccaaa ctgaacaaag
     2041 agctagcttc agctgaacaa aataaaaatc atataaataa tgagctaaag aaaaaggaag
     2101 aacagttgtc cagttatgaa gataagctat ttgatgtttg tggtagccaa gatttggaaa
     2161 gtgatttggg cagacttaaa gaagaaattg aaaagtcctc aaagcagcga gccatgctgg
     2221 caggagccac agcagtttac tcccagttca ttacgcagct gacagatgaa aaccagtcct
     2281 gctgcccagt ctgtcagaga gtatttcaga cagaggctga gttacaggaa gtcatcagcg
     2341 acttgcagtc caagctgagg cttgctccag ataaactcaa gtcaacagaa tcggaactga
     2401 agaaaaaaga gcggcgccgt gatgaaatgc ttgggcttgt gcccgtgagg caaagcataa
     2461 ttgatttgaa ggaaaaggaa ataccagaat taagaaaccg actacagagt gtcaatagag
     2521 acatacagcg cctaaagaat gacatagaag agcaggaaac actcttgggt acaataatgc
     2581 ctgaagaaga aagtgctaaa gtgtgcctga cagatgttac aatcatggag aggttccaga
     2641 tggagctgaa agatgttgaa aggaaaattg cacagcaggc agctaagctg cagggagtag
     2701 acttggaccg gactgtccag caggttaacc aggaaaagca ggaaaagcag cacaggctgg
     2761 atacagttac cagtaagatt gaattgaacc gtaagcttat acaggaccag caggaacaag
     2821 tccaacacct gaaaagtaaa acaaatgagc tgaagtcaga gaagctgcag atagccacca
     2881 atttgcaacg gcgccagcaa atggaggagc agagtgtgga attatccact gaagtccagt
```

```
2941 ctttaaacag agaaat●●g gatgctaaag aacaaataag cccttt    g acagcattgg
3001 aaaagttgca gcaggaaaaa gaagaattaa tccacagaaa acatacaagt aacaaaatgg
3061 ctcaggataa aattaatgat atcaaagaga aagttaaaaa tattcatggt tacatgaaag
3121 acatagagaa ttatattcaa gatggaaaag atgactacaa gaagcaaaaa gaaactgaac
3181 ttaatggagt tgctgttcaa ctaaatgaat gtgagaaaca cagagaaaag ataaataaag
3241 acatgggaac catgaggcaa gacatcgaca cacagaagat acaggaaagg tggctacagg
3301 ataaccttac actaagaaaa aggagagatg aactaaaaga agtggaagaa gaaccaaaac
3361 aacacttgaa ggagatggga caaatgcaag ttttacaaat gaaaaatgag catcagaagt
3421 tggaagagaa catagacact ataaagagaa atcacagttt ggcattgggg cgtcagaaag
3481 gctatgaaga cgaaattctt cactttaaaa aggagctccg agaacctcag ttccgggatg
3541 ccgaggaaaa gtacagagaa atgatgatcg tcatgaggac cacagagctg gtgaacaagg
3601 acctcgacat ctactacaag actcttgacc aagcaattat gaaatttcac agtatgaaaa
3661 tggaagaaat caacaaaatt attcgtgatc tttggcggag tacctatcgt gggcaagata
3721 ttgaatacat agaaattcga tctgatgctg atgaaaatgt atcggcttct gataaaaggc
3781 gaaattacaa ctaccgagtg gtgatgctga aggggacac agccttggac atgcgtggac
3841 gatgcagtgc tggacagaag gtgctagcgt ctctcatcat ccgactggcc ctggcggaaa
3901 ccttctgtct gaactgtggc atccttgcct tggatgagcc gacaacaaat ctcgaccgag
3961 aaaacattga gtctcttgca catgctttgg ttgagatcat aaaaagtcgc tcacagcagc
4021 gcaacttcca acttctggta atcactcacg atgaagattt tgtggagctc ttagggcgat
4081 ctgagtatgt ggagaaattc tacagggtga agaagaacat ggaccagtgc tcagagattg
4141 tcaagtgcag catcagctcc ctgggttctt atgttcacta acttgtcttg ggaaggatat
4201 ggattgtgtg acgggaacca tgtgtgtcag gtttgagagt ccctaagatc gggagggtgc
4261 cttgagtgaa gatgaacaca gttcccacag gcctagcaac ctcccagcac cactcagcct
4321 ttcatttcct gccaactgaa tgtcatccag ataaacagtg ctgacaggag gggtggacat
4381 ttaaacacat tactattttc tctgaagctt gaaggtatat ttaatacgta gccctgctgt
4441 cacatgaatc catgcacacg acctgtctga gtctaggtct cccctgaaac cacagtcatt
4501 gaacggggct ggcagagtga cagaagccat cggtgatcat ttagtggaat acgcagttcg
4561 ctccccatgc agccatacca gataggtacc ctggctggct gttagagaag agttccccca
4621 gtcctgacct catctgccca tcatgcagaa gcctggcgct ctcctctgtg tctgtcatgt
4681 ctatctatgt ggcctggccc agttactctc aatagcaggt tctcctgtgg ctggaaagga
4741 gactcacgca gctggagcct gctccacgca cagacctgca aaggcaaacc acattttcac
4801 tccagaggct gagcctggtg aagtccagta cacatctcca cggcattgca gctgaagcct
4861 cctgtgaaga caggacacat ctaagtcaca cgcttccaga tgtcactaac agttctctca
4921 ctcagctggg tggcctgttc tgtgccatct gcagatttgt tgggtgtcac ctctagaaca
4981 aaggtttttt cttcccactg tcttccttca ggaatataaa accacgtcaa tcaaatgcat
5041 actgagaata gatcataaga aaaattactt taaaataaaa aaaaaaaa
//
```

Revised: October 24, 2001.

APPENDIX A

```
!!AA_MULTIPLE_ALIGNMENT 1.0 PileUP
Symbol comparison table: genrundata:blosum62.cmp   CompCheck: 1102
GapWeight: 8     GapLengthWeight: 2
MSF: 1380   Type: P   February 11, 2002 15:38   Check: 219 ..
```

```
        U63139aa        Human Rad50
        Z75311aa        Human Rad50
        NM022246aa      Rat Rad50
        U66887aa        Mouse Rad50
        AAD15407aa      Arabidopsis Rad50
        AF168748aa      Arabidopsis Rad50
        1116sid2        SEQ ID NO: 2 Maize Rad50
        X14814aa        Yeast Rad50
        Z75312aa        C. elegans Rad50
```

FORMATTING: Identical residue

```
                1                                                  50
   U63139aa     ~~~~~~~MSR  KMSILG RS FGI DKDK   LTFFSPLT  VGPNGAGKTT
   Z75311aa     MLIFSVRDMF AKMSILG RS FGI DKDK   TLFFSPLT  VGPNGAGKTT
 NM022246aa     ~~~~~~~MSR  KMSILG RS FGI DKDK   I FFSPLT  VGPNGAGKTT
   U66887aa     ~~~~~~~MSR  KMSILG RS FGI DKDK   I FFSPLT  VGPNGAGKTT
 AAD15407aa     -------mstv dkmlikgirs r dp nknv  fflp pltli vgangagktt
 AF168748aa     ------MSTV DKMLIKGIRS F DP NKNV  IFFP PLTLI VGANGAGKTT
   1116sid2     ~~~~~~~MSTV DKMLIKGIRS F DPDNKNV  ITFFKPLTLI VGPNGAGKTT
   X14814aa     ~~~~~~~MS  YK SIQGIRS E DSND ET  TEFGKPLTLI VGMNG GKTT
   Z75312aa     ~~~~~~~M KF L HI GIRS VGD DHDVHK IDELSPCTLI SGPNG GKTI

                51                                                 100
   U63139aa     TLECLKYIC  G FPPG  .G NTFVHDPKVA QET V AQI  LQF DVNGEL
   Z75311aa     TLECLKYIC  G FPPG  .G NTFVHDPKVA QET V AQI  LQF DVNGEL
 NM022246aa     TLECLKYIC  G FPPG  .G NTEVHDPKVA QET V AQI  LQF DVNGEM
   U66887aa     TLECLKYIC  G FPPG  .G NTEVHDPKVA QEI V AQI  LQF DVNGEM
 AAD15407aa     lleclk sct gelppn  sg h f hdpkva getetkaqik lrfktaagkd
 AF168748aa     LIECLK SCT GELPPN RSG H F HDPKVA GETETKAQIK LRFKTAAGKD
   1116sid2     LIECLKLSCT GELPPNSRSG HTFVHDPKVA GETETKGQIK LRFKTAAGKD
   X14814aa     TIEQLRY TT G LPPNS GG .VF HDPK  GEK I AQ K LAFT ANGLN
   Z75312aa     TLEALNFVT  GQ P.. QKK QNF HSTDVA RKTRVDAS T LEFIDVKG

                101                                                150
   U63139aa     AVQRSMVCT OK KKTEFK  EGV .IRTK H GEKVSLSS  CAI DREF I
   Z75311aa     AVQRSMVCT OK KKTEFK  EGV .IRTK H GEKVSLSS  CAI DREF I
 NM022246aa     VLVQRSMLC OK KKTEFK  EGV .IRIK H GEKVSLSS  CAI DREF I
   U66887aa     VAVHRSMLC QK NKKTEFK EGV .IRMK H GEKVSLSS  CAI DREF I
 AAD15407aa     vvclrsfdlt qkaskme ka iesvlqtinp htgekvclsy rcadmdreip
 AF168748aa     VVCLRSFQLT QKASKME KA IESVLQTINP HTGEKVCLSY RCADMDREIP
   1116sid2     VVCLRSFQLT QKASKMEFKA IESVLQTINP HTGEKVCLSY RCADMDREIP
   X14814aa     VTRNIQLL MKK TTTEFK  EGQLV IN. N G STLST RSL DAQ P
   Z75312aa     CTA RRLV T .SG RAAALA EEHTLAIKYP .DGTVNTLSS VCDFNTA L
```

```
            151                                                           200
U63139aa    S GVSKAVL NNVIFCHQ  SNWPLS GK  LKQKFD IFS ATRYIKALET
Z75311aa    S GVSKAVL NNVIFCHQ  SNWPLS GK  LKQKFD IFS ATRYIKALET
NM022246aa  C GVSK VL NNVIFCHQ  SNWPLS GK  LKQKFD IFS ATRYIKAL T
U66887aa    C GVSK    NNVIFCHQ  SNWPLS GK  LKQKFD IFS ATRYIKAL T
AAD15407aa  almgvskal  epvlfvhqde snwplqdpst lkkkiddifs atrytkalev
AF168748aa  ALMGVSKAL L ENVIFVHQDE SNWPLQDPST LKKKFDDIFS ATRYTKALEV
1116sid2    ALMGVSKAVF ENVIFVHQDE SNWPLQDPST LKKKFDDIFS ATRYTKALEV
X14814aa    LY GVPKAL L EYVIFCHQ  SLWPLS PSN LKKKE IFQ AM TKALEN
Z75312aa    KH GVP AVF KYVIFCHQ  STWPLS PKE LKKK DDIRQ LT VKAQER

            201                                                           250
U63139aa     Q RQTQGQ K KEYQ ELK YLKQYK KAC E R I SKE AQLT SKEI
Z75311aa     Q RQTQGQ K EEYQ ELK YLKQYK KAC E R I SKE AQLT SKEI
NM022246aa   Q RQTQGQ K KECQTELK YLRQNK KAC E R I SKE AQLA S EI
U66887aa     Q RQTQGQ K KECQTELK YLKQNK KAC E R I SKE AQLA SQEI
AAD15407aa   lkklhtsqaq elkff lkle nlgt kdaay klresiaqdq e    skvqm
AF168748aa   TKKLHKDQAQ ELKTF LKLE NLQT KDAAY KLRESIAQDQ E    SKVQM
1116sid2     TKKLHKDQMQ EIKTF LKLE NLQT KDQAH KLRENIAQDQ EKD  SKSQM
X14814aa     KS KKDMSV LKLL QS E HLKLDKDR K A  LNIHQLQ TKIDQYNEE
Z75312aa     KK VLDFKK E QTHE SKQ LYE HVRDKL VARQNQEECE R ISKRKEET

            251                                                           300
U63139aa    KSYENE DPL KNR KE EHN LSK M LDNE KALD R KQ EKD SELEE
Z75311aa    KSYENE DPL KNR KE EHN LSK M LDNE KALD R KQ EKD SELEE
NM022246aa  KAYENE EPL KNR KE EHN LSK MRLDNE KALD R KQ EKD S ELEQ
U66887aa    RSYE E EPL KNR KE EHN LSK M LDNE KALE R KQ EKD S ELEQ
AAD15407aa  le ets qkv ae h kemm k n iqdq siktberst lfkeqqrqya
AF168748aa  LE ETS QKV AE H KEMM LK ER LQDQ SIKT ERST LFKEQQRQYA
1116sid2     QLKE CGT EDL METS DDELRRLQGQ I IKA ERST LLTQQHEKLA
X14814aa    SE ESQ NEI TEKSDK FKS NQ FQ LSK NLKNT L SDQVKRLSN
Z75312aa    E KE KANG Q KIEEMRIS HELEDTLTS FKKTELERQN LKKQLSLIRV

            301                                                           350
U63139aa    K EKVFQG D EQLN LYHNH QRT .REKER VDCHR  KLNKESR LN
Z75311aa    K EKVFQG D EQLN LYHNH QRT .REKER VDCHR  KLNKESR LN
NM022246aa  K EKVFQG D EQLN LYHNH QRT .REKER VDCQR  KL KE R LN
U66887aa    K EKVFQG D EQLN LYHNH QRT .REKER VDCQR  KLNKE R LN
AAD15407aa  alpeenedri eelkewk kf eer .allgt krk eranv dte tisslh
AF168748aa  A PEENEDTI EELKEWK KF EER .ALLGT KIRK ER MV DTE TISSLH
1116sid2    ALSEENEDTD EELMEWQ KF EER .ALLET KI KLVRDM DE YSS LS
X14814aa     D.IL L K P QNLL NE SKV .MDKNN Q RDLETD S SLKDRQSSLQ
Z75312aa    ...EPYFG  EELKREIEE RGSEGRSYGE ER Q K G KNNQERQELS
```

2

```
             351                                                           400
U63139aa     QEKSELLVEQ GLLQLQADRH QEHIRARDSL IQSAATQLE. LDGFERGPF
Z75311aa     QEKSELLVEQ GLLQLQADRH QEHIRARDSL IQSAATQLE. LDGFERGPF
NM022246aa   QERAELLVEQ GLLQLQADRH QEHIRARDSL LQSAAHLE.  LDGFERGPF
U66887aa     QEKAELLVEQ GLLQLQADRH QEHIRARDSL IQSAATHLE. LDGFERGPF
AAD15407aa   nakanymlek skiqteadah mlaknerdst lqniffhyn. lgnvpstpf
AF168748aa   NAKANYMLEK SKLQTEADAH MLAKNERDST LQNIFFHYN. LGNVPSTPF
1116sid2     KQNSELTHRI GKLQAEADAH IAKHERDSD IKNICTKHN. LGPVEEHPF
X14814aa     SLSNSIIRRQ GEGEAGKEIY EKNANHLSSL KEAFQHKFSG LSNAENSDM
Z75312aa     QKKAEFENRI SSLKAEVIHC QSAKYALARL ENARREELD. ...AEHDADI

             401                                                           450
U63139aa     ERQIKNFHKL AERQEGE.A KTANQLMNDF AAKETLKQKQ ID.......E
Z75311aa     ERQIKNFHKL AERQEGE.A KTANQLMNDF AAKETLKQKQ ID.......E
NM022246aa   ERQIKNFHEL AERQERE.A KTASQLLSDE TDKEALKQRQ ID.......E
U66887aa     ERQIKNFHEL RERQERE.A KTASQLLSDE TDKEALKQRQ ID.......E
AAD15407aa   tavvanltnk lktb          lgelemall ldkkksneta lstawdcyad
AF168748aa   TAVVNLTNR IKLR          LGELEMDL LDKKKSNETA LSTAWDCYAD
1116sid2     NAVANLTNR IKAR           ...LSSLENDL LDKKKSNEDQ LDVAWKHYAK
X14814aa     ..QVNHEMSQ FKAFISQDLT DTADQFAKLA QLKETNLSDL AKSATVDSQN
Z75312aa     DIEIDNAIT. AKIRGMSDKA RMANKNCALL QSNLRAAQAA ATKAEVEMKT

             451                                                           500
U63139aa     ARDAKAGAGR IAE.LKSAI S..KAQNELK NVKYALQQLE GSSDRALEAD
Z75311aa     ARDAKAGAGR IAE.LKSAI S..KAQNELK NVKYALQQLE GSSDRALEAD
NM022246aa   ARDAKSGAGR MAE.LKTAI A..KAQTELA NVRNALQQLE GSSDRALEAD
U66887aa     ARDAKSGAGR TAE.LKTAI A..KAQSELA HVRSALQQLE GSSDRALEAD
AAD15407aa   andraksaaa qkr.akdaik mgisAaieak eierdafefa stvdakqta
AF168748aa   ANDRAKSAAA QKR.AKDAIK MGISARIEAK AIERDAFEFA ASTVDAKQTA
1116sid2     INARASAAG QIQ.AKIESA SGILARKIK EKERDAAEVA ASKFNLSRAD
X14814aa     AEYNKKARSK LAH.DSEELA EKAKSFASLS TQASLNHELE ..........
Z75312aa     ANEKVKAAK ELEQLKFKIK QGQNATAGMK ALLAKEEALR KSLADLPLAD

             501                                                           550
U63139aa     QELIKAEREA SAAEKNSNVE TLKMEAISLQ NEKADLIRTA KLDQEMAQA
Z75311aa     QELIKAEREA SAAEKNSNVE TLKMEAISLQ NEKADLIRTA KLDQEMAQA
NM022246aa   QELTKAEREA SAAEKNSSIE TLKAELLILQ AEKADLIRNA KLDQEMAQA
U66887aa     QELTKAEREA SAAEKNSSIE TLKAEMSLQ NEKADLIRSA KLDQEMAQA
AAD15407aa   efeaqace erktkqnsee glaskieqka heayslehki ktlnneard
AF168748aa   EREAQAQAE ERKIKQNSER GLASKIEQKQ HEAYSLEHKI KTLNRKAD
1116sid2     EREAHACEA ERKTLALGER DYISIISQKR AEAYSLEKKI KVLLREKD
X14814aa     .NLATYKEK QSWESENIIP KLNQKIEEKA NEAIILEAQI EKFQDRIMKT
Z75312aa     ENAL.TECKA KR.......E KYLKQADILK KKCAEAEKNA EK.DREKAS
```

```
            551                                                      600
  U63139aa   NHH TTRTQ E LTKDKADK D Q RK KSR HSD.E TS L .GYF ....N
  Z75311aa   NHH TTRTQ E LTKDKADK D Q RK KSR HSD.E TS L .GYF ....N
NM022246aa   NHH TTRTQ E LTKDKTDK D Q RK KSR HSD.E TS L .GYF ....N
  U66887aa   NHH TTRTQ E LTKDKTDK D Q RK KSR HSD.E TS L .GYF ....N
AAD15407aa   agd  r .. .........l tr..... de ckd.  gv1  g lppekdm
AF168748aa   AGDA RVKL S KTEQENL KK HKK DE CKD.  GVL  GRLPPEKDM
  1116sid2   NR AD RVKL G KKD SS KDKL EI NE HKD.  KVD RGRNFEKDM
  X14814aa   NQ ADLYAK  G IKK NTK LDEL K TEK LQN.D...SR  RQVF LTQ F
  Z75312aa   KQ..TLS AR KKMTAYQRIY DNNW G GQ AP FPWTP L SKTFHKL ND

            601                                                      650
  U63139aa   KKQ EDWLHS KS EINQTRD LAKLNKE A S EQNKNHIN NE K KEEQ
  Z75311aa   KKQ EDWLHS KS EINQTRD LAKLNKE A S EQNKNHIN NE E KEEQ
NM022246aa   KKQ EDWLHS KS EINQTRD LAKLNKE A S EQNKNHIN NE KKKEEQ
  U66887aa   KKQ EDWLHS KS EINQTRD LAKLNKE A S EQNKNHIN NE KKKEEQ
AAD15407aa   k  c vg lrs   yd sl ksre ske n ml mk q vn ns fkhn  t
AF168748aa   K  V Q LRS   EYD LSL KSREA KE N MLQMK Q VN NSLFKHN  T
  1116sid2   K EINQAFWP V KEYN RS KSQ A QE K F QSK  AR EQLTK R D
  X14814aa   Q  ..ADLEMD FQ L IN QK NIAINNKK H ELDR YINAL Y LNT E D
  Z75312aa   KKI EEDLRD V QLNV..... .......... ...QK ETMQ H YR QEES

            651                                                      700
  U63139aa   S  .... E K LFDVCGSQDF ESD..... D RLKEEI KSS QRAMLA AT
  Z75311aa   S  .... E K LFDVCGSQDF ESD..... D RLKEEI KSS QRAMLA AT
NM022246aa   S  .... E K LFDVCGSQDF ESD..... D RLKEDI KSS QRAMLA AT
  U66887aa   S  .... E K LFDVCGSQDL ESD..... G RLKEEI KSS QRAMLA AT
AAD15407aa      .   e   q  kge v t da pk le sakdkrd rk eynmangm
AF168748aa      . R  S KLQ  KQESV T DA PK LE SAKDKRD RK REYNMANGMR
  1116sid2   R . R  S KLQ  LQISA N DM PK L DA NKRD QK RLENFANGMR
  X14814aa   QDNQ. SK K V L .SENL PE CTIDEY DV EET LSY TALENLKMH
  Z75312aa   T QEL LS N IFE C...SC EA EVSEK E NLRK LKKA DLAPLSAKS

            701                                                      750
  U63139aa   A Y Q ITQL TD..ENQSC  VCQRV QT A LQEV SDL QSKLRL P
  Z75311aa   A Y Q ITQL TD..ENQSC  VCQRV QT A LQEA SDL QSKLRL P
NM022246aa   A Y Q ITQL TD..ENQSC  GCQRV QT A LQEV SDL QSKLRL P
  U66887aa   A Y Q ITQL TD..ENQSC  VCQRV QT A LQEV SDL QSKLRL P
AAD15407aa   q fep ekra q..ehs.. c pc er  ad ee .asf kg  kasstgeh
AF168748aa   QMFEP EKRA RQ..EHS.. C PCGER IAD EE.ASF KQ R KASSTGEH
  1116sid2   EML PFEHLA RK..NHV.. C PCCER PTPD EE.DEF KKQ R QNSSTAK
  X14814aa   QTTLE NRKA LEIAERDSC  YL SRKFEN SFKSKL QEL TKTD NFE
  Z75312aa   N YDSYI... .EESKSSGC  LC RDEKTK K INE SK L ENMTLSFPTE
```

4

```
              751                                                      800
  U63139aa    LK TESELKK KEKRRDE LG LVP RQ I D EKE..KEIPE LRNKL VNR
  Z75311aa    LK TESELKK KEKRRDE LG LAP RQ I D EKE..KEIPE LRNKL VNR
NM022246aa    LK TESELKK KERRRDE LG LVP RQ I D EKE..KEIPE LRNRL SVNR
  U66887aa    LK TESELKK KERRRDE LG LVP RQ I D EKE..KEIPE LRNRI SVNR
AAD15407aa    lkala essn a  fqqldk lr  eysk ltt..eiipl aektl ehte
AF168748aa    LKALA ESSN A  EQQLDK LR V EYSK LTT..EIIPL AEKTL EHTE
  1116sid2    KALA RSSN AE  EQQLDK LR  AY K LVE..ETIPL AEKNL HL
  X14814aa    LKD QNEK .....EYLHS LRL EKHI T LNSINEK DN QKCLEKAKE
  Z75312aa    QEELEKLVSK LEK..EE II  EGQA.NE LQRIVKE KE VREKNRKLS
```

```
              801                                                      850
  U63139aa    DIQRLK DI  QET LGT M P E AKVCL TD.VTI RF QME KD E K
  Z75311aa    DIQRLK DI  QET LGT M P E AKVCL TD.VTI RF QME KD E K
NM022246aa    DIQRLK DI  QET LGTVM P E AKVCL TD.VTI RF QME KD E K
  U66887aa    DIQRLK DI  QET LGT M P E AKVCL TD.VTI RF QME KD E K
AAD15407aa    lgqk eala d lq saq k ad d eal  qp.l nadri fqelvsyqkq
AF168748aa    LGQK EALD D LG SAQ K AD D EAL  QP.L NADRI FQEIVSYQKQ
  1116sid2    DESQK AFD DLLG IAHV MD D EALL QP.T TIDRH VHEIQQLVEE
  X14814aa    TKTSKSKLD EIEVDS KL K D LAE E  RELI KFTYL EKE KDLENS
  Z75312aa    MAEEKSNLS KNEKQLETV  AKLKLAED .L QTDVGVIQQL YEQTEENEKR
```

```
              851                                                      900
  U63139aa    AQQAAKLQG IDLD. QQ  QEK EKQH KLDT SSK  LNRKLIQDQ
  Z75311aa    AQQAAKLQG IDLD. QQ  QEK EKQH KLDT SSK  LNRKLIQDQ
NM022246aa    AQQAAKLQG VDLD. QQ  QEK EKQH KLDT SSK  LNRKLIQDQ
  U66887aa    AQQAAKLQG VDLD. QQ  QEK EKQH RLDT TSK  LNRKLIQDQ
AAD15407aa    edleykfaf rglgvk ee lqselsslgs  dkihge  klrddqiy e
AF168748aa    EDLEYKLDF RGLGVK EE LQSELSSLGS  DKLHGE  KLRDDQIY E
  1116sid2    EDLEYALDS SGRGVK EE LQLEL FLQ  DTLIVE  DLRDQHRM
  X14814aa    SKT SEE SI YNTSEDG QT  DELRDQQR  MND LRELRK T SDLQMEKD
  Z75312aa    YEQLVSESDS SDGLSY ... ... ELRKKVE D DEEYRK V QEG ELQKCS
```

```
              901                                                      950
  U63139aa    EQ QHLKSTT NEL SEK QI S...N QRR QQLEEQTVEL TEVQSHYRE
  Z75311aa    EQ QHLKSTT NEL SEK QI S...N QRR QQLEEQTVEL TEVQSHYRE
NM022246aa    EQ QHLKS T NEL SEK QI ...N QRR QQMEEQTVEL TEVQSLNRE
  U66887aa    EQ QHLKS T NEL SEK QI ...N QRR QQMEEQSVEL TEVQSLNRE
AAD15407aa    rd scl  rw havreekaka  ...n lrdv tk ee lerl aeeksold
AF168748aa    RD SCLQARW HAVREEKAKA  ...N LRDV TK EE LERL AEEKSOLD
  1116sid2    ED SSAQVRW HNAREEKYKA S...S LERF QK EKLVLL AEEKEQIIE
  X14814aa    EK RENSRMI NLI EKE TV SEIESSHTQK QNI S RSK RENIND DSR
  Z75312aa    E RNKLQS L NELGTH VSL GEAA ...QA GAFAEQLETK IKEIQECITA
```

5

```
            951                                              1000
   U63139aa    IKDAK.EQ S  PLETTLEKFQ  QEKEELIKKK  NTSNKIAQ  LN IKEK KN
   Z75311aa    IKDAK.EQ S  PLETTLEKFQ  QEKEELIKKK  NTSNKIAQ  LN IKEK KN
 NM022246aa    IKDAK.EQ N  PLEIALEKLQ  QEKEELIH K  NTSNKMAQ  IN IKEK KN
   U66887aa    IKDAK.EQ S  PLETALEKLQ  QEKEELIH K  HTSNKMAQ  IN IKEK KN
 AAD15407aa    vkylt.e lg  plskekeqll  s ynd ki r  nqeyeelae  kin gge a
 AF168748aa    VKYLT.E LG  PLSKEKEQLL  S YNDMKI R  NQEYEEIAE  KRN QQE A
   1116sid2    KKLLE SID  PLSKEKESLL  QEYN LK RL  DEEYHELAE  KRE QQELDA
   X14814aa    VKELE AR I  SLKNKK EAQ  SVLDK K ER  DIQVR KQKT  VA IRL DR
   Z75312aa    ISQKRNED P  DAQFKK D T  RNVS KEEEK  KKAEME QMM  K SLDQK FH

            1001                                             1050
   U63139aa    HGYMKD EN  YIQDGK DYK  KQ.KETEL K  ....VI Q S  E EKHKE N
   Z75311aa    HGYMKD EN  HIQDGK DYM  KQ.KETEL K  ....VI Q S  E EKHKE N
 NM022246aa    HGYMKD EN  YIQDGK DYK  KQ.KETEL E  ....VVIQ   E DKHKE N
   U66887aa    HGYMKD EN  YIQDGK DYK  KQ.KETEL Q  ....VAVQ   E EKHRE N
 AAD15407aa    ill asyk nd  cftr hdlkk  ge. d ge  qrlsdsqlq  scea k e
 AF168748aa    ILL ASYK NE  ....SHDLKK  GE. ED QE  QRLSDSQLQ  SCEA K E
   1116sid2    LIG LNMK K.  ...GYLDSKK  NE.KLKEL QG  HVL CHSQLQ  SCMA Q RIS
   X14814aa    FQT YNE VD  FEAKGFDELQ  TTIKELRI K  ...... Q L  ELKEQLDLKS
   Z75312aa    RKS FKK QE  ...GGLCE Q  LMDKENN ..  ...ATLN SLE  ENQQ QKRFE

            1051                                             1100
   U63139aa    E RLM QD  DTQKIQERW  QDNLTLRKRN  E KE E...  EERKQHLKE
   Z75311aa    E RLM QD  DTQKIQERW  QDNLTLRKRN  E KE E...  EEGKQHLKE
 NM022246aa    KE GTM QD  DTKKIQERW  QDNLTLRKR  E KE E...  EERKQHLKE
   U66887aa    K GTM QD  DTQKIQERW  QDNLTLRKR  D KE E...  EEPKQHLKE
 AAD15407aa    ge n nk l  rngd l mi  onlny tk  akveeltr  esleeq lni
 AF168748aa    GELN NK L  RNQDQ RN  DNLNYRTK  AKVEELIR I  ESLEEQ LNE
   1116sid2    AELN SK LL  QQGQLARN  DNLKYRKTK  A VEQLTR L  ESLEERLLSI
   X14814aa    NEVNEEKRK  ADSNNEE N  KQNLELIELK  Q QH ES I  SRL... QN
   Z75312aa    E RS....F  DSSH RESI  KD LTRMIIE  NK KELKRT  A F GQ NED

            1101                                             1150
   U63139aa    GQ QV QMK  HQ LEENI  N KRNHNLAL  GRQKGYEEE  IHFKKEL EP
   Z75311aa    GQ QV QMK  HQ LEENI  N KRNHNLAL  GRQKGYEEE  IHFKKEL EP
 NM022246aa    GQ QV QMKN  HQ LEENI  T KRNHSLAL  QRQKGYEEE  LHFKKEL EP
   U66887aa    GQ QV QMKN  HQ LEENI  T KRNHSLAL  QRQKGYEDE  LHFKKEL EP
 AAD15407aa    gg ale.. a  v ilre e  r llselntcr  gt svyessi  skn velkq
 AF168748aa    GG ALE.. A  V ILRE E  RLLSELNRCR  GT SVYESSI  SKN VELKQ
   1116sid2    GS ALE.. A  KRHS E E  RLNSEFNRWQ  GT SVYQS I  SKHQELKL
   X14814aa    AEAERDKYQE  ESLRLRTRFE  SSENAGKL  GE KQLQN I  DSLTHQL .
   Z75312aa    ...RITEQKQ  AYN LQ ELR  L GNEEV IY  TQ QEYEK K  KIAEAKLSTK
```

6

```
              1151                                                      1200
   U63139aa    Q..DAEE.YR  EMM.V..TE  VNKDLDIYY  K.LDQA.M.F  HSMKMEEINK
   Z75311aa    Q..DAEE.YR  EMM.V..TE  VNKDLDIYY  K.LDQA.M.F  HSMKMEEINK
  NM022246aa   Q..DAEE.YR  EMM.V..TE  VNKDLDIYY  K.LDHA.M.F  HSMKMEEINK
   U66887aa    Q..DAEE.YR  EMM.V..TE  VNKDLDIYY  K.LDQA.M.F  HSMKMEEINK
  AAD15407aa   qykdi.krhf  dql.ciktts mankdldryy  naldkalmrf  h.mqmeeink
  AF168748aa   QYKDI.KRHF  DQL.OLKITE MANKDLDRYY  NALDKALMRF  H.MKMEEINK
    1116sid2   QYKDIEKRYT  NQ..OLKITE MANKDLDRYY  TALDKALMRF  HSMKMEEINK
   X14814aa    DYKDIEKNYH  KE..ELQIRS FVTDP.DVYS  KALDSA.M..  HG.KMQ.INK
   Z75312aa    ECQNAESNYR  DAI.ELAITK E.ISDIT.YR  NCLDA.L.QF  HSEKMGR.NG

              1201                                                      1250
   U63139aa    .I..LWRSTY  RGQDI.YIEL  RSDADENVSA  SDKRRNYNYR  VVMLKGD.AL
   Z75311aa    .I..LWRSTY  RGQDI.YIEL  RSDADENVSA  SDKRRNYNYR  VVMLKGD.AL
  NM022246aa   .I..LWRSTY  RGQDI.YIEL  RSDADENVSA  SDKRRNYNYR  VVMLKGD.AL
   U66887aa    .I..LWRSTY  RGQDI.YIEL  RSDADENVSA  SDKRRNYNYR  VVMLKGD.AL
  AAD15407aa   .i..elwqqty rqgd.dyirl  hsd.....se  gagtrsysyk  v.mqtgd.el
  AF168748aa   .I..ELWQQTY RGQD.DVIRI  HSD.....SE  GAGTRSYSYK  VVMQTGD.EL
    1116sid2   .I..ELWQQTY RGQD.DYISI  NSD.....SE  GAGTRSYSYR  VVMQTGD.EL
   X14814aa    .I.DELWKRIY  SGTD.DTIKI  RSD...EVSS  TVKG.SYNYR  VVMYKQDVEL
   Z75312aa    .IID.LWRKVY  NSTDITT.RI  RSDATSETSS  KKV.YEYNVM  .VHETG..EI

              1251                                                      1300
   U63139aa    MRGRCSAGQ.  KVLASLI.RI  ALAETFCLNC  GIIALDEPTI  NLDRENIESL
   Z75311aa    MRGRCSAGQ.  KVLASLI.RI  ALAETFCLNC  GIIALDEPTI  NLDRENIESL
  NM022246aa   MRGRCSAGQ.  KVLASLI.RI  ALAETFCLNC  GIIALDEPTI  NLDRENIESL
   U66887aa    MRGRCSAGQ.  KVLASLI.RI  ALAETFCLNC  GIIALDEPTI  NLDRENIESL
  AAD15407aa   emrgrcsagq  kvlaslilti  alaetfclnc  giialdeptt  nldgpn.esl
  AF168748aa   EMRGRCSAGQ  KVLASLILRI  ALAETFCLNC  GIIALDEPTT  NLDGEN.ESI
    1116sid2   EMRGRCSAGQ  KVLASLIIRI  ALAETFCLNC  GIIALDEPTT  NLDGPNAESL
   X14814aa    MRGRCSAGQ.  KVLASL.ITRI AI.ETFGANC  GI.ALDEPTI  NIBEENIESL
   Z75312aa    EMRGRCSAGQ  K.LASL.IRI  ALAEV.GGSC  S.IALDEPTI  NLDESKVEG.

              1301                                                      1350
   U63139aa    AHAT.ET.K.  RSQ........  .QRNFQL.VI  THDEDFVEL.  GRSEYVEK.Y
   Z75311aa    AHAT.ET.K.  RSQ........  .QRNFQL.VI  THDEDFVEL.  GRSEYVEK.Y
  NM022246aa   AHAT.ET.K.  RSQ........  .QRNFQL.VI  THDEDFVEL.  GRSEYVEK.Y
   U66887aa    AHAT.ET.K.  RSQ........  .QRNFQL.VI  THDEDFVEL.  GRSEYVEK.Y
  AAD15407aa   aga.lrimed  rkg........  .qenfqlivi  chderfaq.i  garghaekyy
  AF168748aa   AGA.LRIMED  RKG........  .QENFQLIVI  THDERFAQ.I  GQRQHAEKYY
    1116sid2   AAA.LRIME.  RKG........  .QENFQLIVI  THDERFAHLI  GQRQLAEKYY
   X14814aa    AK..LHNI.NM  R.H........  .QKNFQLIVI  THDEK.LGH.  NAAAF..H.
   Z75312aa    AIVLADI.AE  R.GFDENGKL  RGRDMO..VI  THDERLVNRI  TISCRPEYIY
```

7